

Acta Linguistica Hungarica

AN INTERNATIONAL JOURNAL OF LINGUISTICS

Volume 47 Numbers 1–4
2000



Akadémiai Kiadó
Budapest



Kluwer Academic Publishers
Dordrecht/Boston/London

ACTA LINGUISTICA HUNGARICA

AN INTERNATIONAL JOURNAL OF LINGUISTICS

Acta Linguistica Hungarica publishes papers on general linguistics with particular emphasis on discussions of theoretical issues concerning Hungarian and other Finno-Ugric languages. Papers presenting empirical material must have strong theoretical implications. The scope of the journal is not restricted to the traditional research areas of linguistics (phonology, syntax and semantics, both synchronic and diachronic), it also covers other areas such as socio- and psycholinguistics, neurolinguistics, discourse analysis, the philosophy of language, language typology, and formal semantics.

Book and dissertation reviews are also published. The articles should be written in English, exceptionally, however, papers written in German and French are also accepted.

Manuscripts and editorial correspondence should be addressed to:

ACTA LINGUISTICA HUNGARICA
Institute for Linguistics
H-1399 Budapest, P.O. Box 701/518
Phone: (36 1) 351 0413
Fax: (36 1) 322 9297
E-mail: kiefer@nytud.hu
siptar@nytud.hu

Distributors:
for Hungary

AKADÉMIAI KIADÓ
P.O. Box 245, H-1519 Budapest, Hungary
Fax: (36 1) 464 8297
Homepage: www.akkrt.hu

for all other countries

KLUWER ACADEMIC PUBLISHERS
P.O. Box 17, 3300 Dordrecht, The Netherlands
Fax: (31) 78 639 2254
Homepage: www.wkap.nl

Publication programme, 2000: Volume 47 (in 4 issues).
Subscription price: NLG 600.00 (USD 273.00)
per annum including postage & handling.

© Akadémiai Kiadó, Budapest 2000

Acta Linguistica Hungarica is abstracted/indexed in Current Contents—Arts and Humanities, Arts and Humanities Citation Index and Bibliographie Linguistique/Linguistic Bibliography

ACTA LINGUISTICA HUNGARICA

AN INTERNATIONAL JOURNAL OF LINGUISTICS

MANAGING EDITOR:

FERENC KIEFER

HUNGARIAN ACADEMY OF SCIENCES, BUDAPEST

ASSOCIATE EDITOR:

PÉTER SIPTÁR

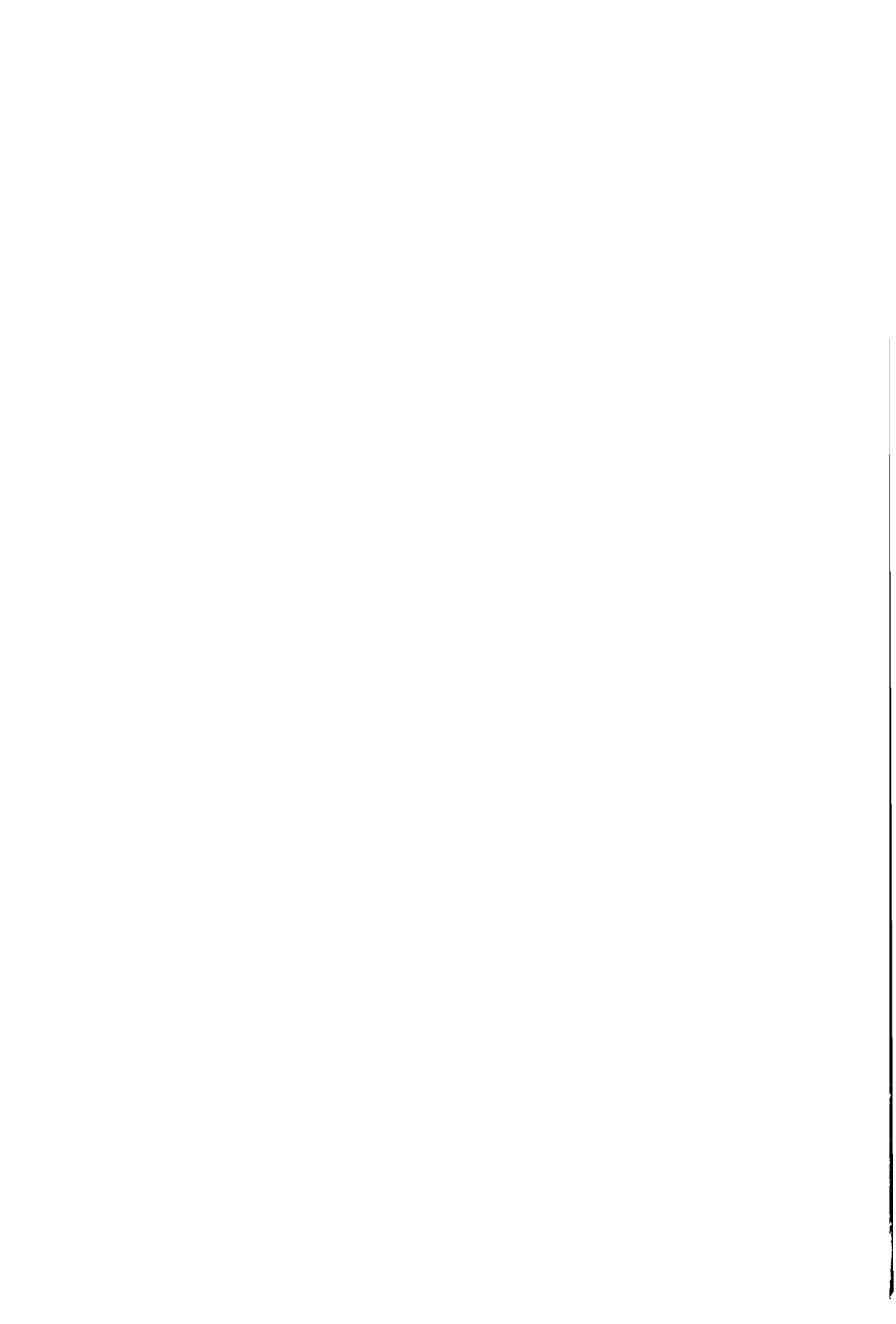
EÖTVÖS LORÁND UNIVERSITY, BUDAPEST

EDITORIAL BOARD:

| | |
|----------------------------|------------------------------------|
| HUBERT HAIDER | UNIVERSITÄT SALZBURG |
| MARTIN HASPELMATH | UNIVERSITY BAMBERG |
| HARRY VAN DER HULST | UNIVERSITY OF CONNECTICUT |
| ISTVÁN KENESEI | JÓZSEF ATTILA UNIVERSITY, SZEGED |
| ANDRÁS KERTÉSZ | KOSSUTH LAJOS UNIVERSITY, DEBRECEN |
| KATALIN É. KISS | HUNGARIAN ACADEMY OF SCIENCES |
| EDITH MORAVCSIK | UNIVERSITY OF WISCONSIN, MILWAUKEE |
| CSABA PLÉH | JÓZSEF ATTILA UNIVERSITY, SZEGED |

AKADÉMIAI KIADÓ
BUDAPEST

KLUWER ACADEMIC PUBLISHERS
DORDRECHT / BOSTON / LONDON



PREFACE

The present volume of *Acta Linguistica Hungarica* contains a selection of papers presented at the *8th International Morphology Meeting* held in Budapest, 12 through 14 June 1998, organized jointly by the Research Institute for Linguistics of the Hungarian Academy of Sciences and the Institute of Linguistics of the University of Vienna. The motto of the conference was “Morphology by itself”; accordingly, the special topics dealt with included the organisation of paradigms, the description of inflectional classes, distributional properties, the problem of morphological categories, aspects of the lexical representation of morphological information and morphological case. However, the conference was open to other morphology-related topics as well. László Elekfi’s paper was not presented at the conference, but it has been included since its topic was found highly relevant to the main theme of the conference. The obituary of Claus-Jürgen Hutterer, who was the editor (1968–1976) and the managing editor (1976–1982) of our journal, concludes the volume.

Budapest, 31 July 2000

Ferenc Kiefer

VP-ELLIPSIS AND VERBAL INFLECTION IN HUNGARIAN*

HUBA BARTOS

Abstract

The paper treats VP-ellipsis data in Hungarian from the perspective of identity requirements between the elided and the antecedent V-form. The examination of the data yields the generalization that agreement marking apparently falls under the notion of sloppy identity, while tense marking shows little variability. Nevertheless it is argued here that the licensing of ellipsis observes the strict identity requirement: elided Agr is not only recoverable from the antecedent V-form in a parallel clause, but, more locally, from the anchor of agreement (subject, object). Likewise, the ellipsis of infinitival marking can be locally licensed by relying on selectional properties. The framework of the analyses is Distributed Morphology, a late insertion theory, whereby ellipsis is regarded neither as deletion, nor as reconstruction, but as non-insertion of phonological material at the level of morphology.

Introduction

The present paper investigates a narrow segment of VP-ellipsis data, with the aim of trying to determine how it correlates with the structure of verbal inflectional morphology in Hungarian, in order to shed some light on the latter, from a (morpho)syntactic point of view. It is absolutely necessary at this point to stress that for reasons of initial organizability, stemming from the fact that this is a primary investigation of the topic and the data, the scope of analysis is artificially narrowed down considerably, in the following respects:

- (i) only those cases will be treated where the VP-ellipsis occurs in the second conjunct of a coordinate structure, i.e., cases of backward ellipsis and ellipsis in subordinating constructions will be ignored;
- (ii) only those cases will be treated where VP-ellipsis at the PF level actually means the ellipsis of the verbal form only, without any of its complements

* I am indebted to Misi Bródy for discussions, Katalin É. Kiss and Richard Kayne for comments, Zoltán Bánréti for helpful suggestions, and the audiences of the *1st Szeged Linguistics Conference for Doctoral Students* (November 1997), and the *7th International Morphology Meeting* (Budapest, June 1998) for comments and criticism. The research has been part of a project supported by OTKA (National Scientific Research Fund), under grant no. T 018131.

or other modifiers, i.e., the examples will be constructed in such a way that at the spellout point the VP will only contain V, either because it lacks any modifiers, or if it has some, they will have left the VP (e.g. by focusing or topicalization)

The treatment of such further data is simply left for future research, with the tacit assumption that the analysis to be offered here is basically suitable for incorporating these in a simple way. At present, however, we are only interested in the conditions imposed on the ellipsis by the morphological make-up of the verb, leaving aside licensing factors pertaining to the properties of other elements in the VP.

The framework of investigation will in its essential aspects be the Minimalist Program of Chomsky (1995), but (as the actual analysis will suggest) we will have to depart from it in several respects. Also, a brief discussion is in order at this point concerning recent construals of (VP-)ellipsis phenomena. While the more traditional view holds that ellipsis is literally deletion of material, i.e., some previously present (group of) lexical item(s) is eliminated from a structure, current generative theories appear to favor the reverse approach, that “ellipsis” in fact means that the items in question are represented in the syntactic structure by some empty element first, which is filled by content only later during the derivation, most plausibly at (or right before) the level of LF, cf. Fiengo–May (1994), Hornstein (1995), whose crucial motivation for this view comes from a particular type of VP-ellipsis called **antecedent contained deletion (ACD)**. Since ACD will not concern me here, and since I believe that my proposal can in principle be extended to cover the data which the LF-reconstruction theories aim to account for, I will assume the “classic” deletion model without any justification, for the ease of exposition. Some discussion of the issue can be found in the Appendix of this paper.

The layout of the paper is the following: Section 1 presents the fundamental Hungarian data (limited in scope as described above). Section 2 presents Lasnik’s (1995) account of VP-deletion phenomena in English, which keeps to the assumption (once held quite widely, in conformity to Sag’s (1976) suggestion) that syntactic deletion observes the criterion of **strict identity** with some antecedent (licensing) form. Section 3 shows that a Lasnik-type analysis is not available for our data, so in Section 4 an alternative is proposed, building on the notion of **late lexical insertion**, cf. Halle–Marantz (1993), Ackema (1995).

1. The key data

It is generally, and quite naturally, assumed that the ellipsis (i.e., the non-pronunciation) of any material from a linguistic form is licit only if the deleted material is recoverable, relying on the linguistic (and sometimes maybe also the non-linguistic) context. In other words, we may choose to omit some part of a form only as long as the omitted information can be reconstructed by the addressee of the communicational act—otherwise this act functionally fails. This state of affairs indicates that acceptability differences in this area can possibly be accounted for in “low-level” spheres of communication, e.g. pragmatics. There are reasons, however, to assume that the explanation for the differences stems from some deeper level, claimed here to be the realm of syntax. In what follows, I will attempt to demonstrate that the syntax-morphology interface renders it not just possible, but also desirable, to make grammaticality distinctions between the forms evaluated differently in acceptability judgements. At the same time, important consequences can be drawn from these judgements as to the morphosyntactic structuring of Hungarian verb forms, hence as to the phrase structure domains of Hungarian, assuming some mirroring effects between morphological structure and syntactic structure, cf. Baker (1985); see also Halle–Marantz (1993).

Let us get acquainted with the fundamental data:

- (1) (a) Péter *alszik* és Maci Laci is [_{VP} ~~*alszik*~~
 Peter sleep-pres-3sg and Yogi Bear also sleep-pres-3sg
 ‘Peter is asleep, and so is Yogi Bear.’
- (b) Péter *alszik* és én is [_{VP} ~~*alszom*~~
 Peter sleep-pres-3sg and I also sleep-pres-1sg
 ‘Peter is asleep, and so am I.’
- (c) Péter *alszik* és Maci Laci is fog [_{VP} ~~*aludni*~~
 Peter sleep-pres-3sg and Yogi Bear also will-3sg sleep-inf
 ‘Peter is asleep, and so will be Yogi Bear.’
- (d) Péter *alszik* és én is fogok [_{VP} ~~*aludni*~~
 Peter sleep-pres-3sg and I also will-1sg sleep-inf
 ‘Peter is asleep, and so will be I.’
- (e) Péter *alszik* de nekem nem kell [_{VP} ~~*aludnom*~~
 Peter sleep-pres-3sg but to-me not need-3sg sleep-inf-1sg
 ‘Peter is asleep, but I don’t have to be.’

As the data show, a verbal item formally identical and functionally–structurally parallel to another one in the preceding conjunct can be deleted: (1a); but so can

non-fully-identical forms: (1b–e). In particular: the licenser and the licensee can differ in ϕ -features (number/person); and an infinitive (whether lacking or displaying overt agreement) can undergo ellipsis after a tense-marked antecedent. Observationally, the conclusion seems to be that **sloppy** (rather than strict) **identity** is at play in the formal licensing of VP-ellipsis—as far as the verb form is concerned. Things, however, are not exactly that straightforward, as shown in (2):

- (2) (a) *Péter ma *táncol*, én pedig tegnap [VP ~~táncoltam~~]
 Peter today dance-pres-3sg I however yesterday dance-past-1sg
 ‘Peter is going to dance today, while I did (= danced) yesterday.’
- (b) Péter tegnap *táncolt*, én pedig ma [VP ~~táncoltam~~]
 Peter yesterday dance-past-3sg I however today dance-past-1sg
 / *[VP ~~táncolok~~]
 dance-pres-1sg
 ‘Peter danced yesterday, while I did so / *am doing so today.’
- (c) *Péter holnap fog *táncolni*, én pedig tegnap [VP ~~táncoltam~~]
 Peter tomorrow will-3sg dance-inf I however yesterday dance-past-1sg
 ‘Peter will dance tomorrow, while I did so yesterday.’
- (d) Péter tegnap *táncolt*, én pedig holnap fogok [VP ~~táncolni~~]
 Peter yesterday dance-past-3sg I however tomorrow will-1sg dance-inf
 ‘Peter danced yesterday, while I will do so tomorrow.’

Clearly, then, sloppy identity does not extend to tense-marking. If the antecedent and the elided V differ in tense-marking, the ellipsis is not properly licensed: (2a, b)—unless the elided form is an infinitive: (2d). On the other hand, an infinitive (non-tense-marked) antecedent does not license the dropping of a tense-marked form: (2c).

It is interesting to note that in many cases of subordination the same paradigm appears. (3) contains illustration with ACD data:

- (3) (a) *Találkoztam* azokkal a fiúkkal, akikkel te is [VP ~~találkoztál~~]
 meet-past-1sg with those boys who-with you also meet-past-2sg
 ‘I met the boys you did, too.’
- (b) *Holnap is *találkozom* azokkal a fiúkkal akikkel tegnap
 tomorrow also meet-pres-1sg with those boys who-with yesterday
 [VP ~~találkoztam~~]
 meet-past-1sg
 ‘Tomorrow I will meet the boys I did (meet) yesterday, too.’

- (c) Tegnap már *találkoztam* a fiúkkal akikkel ma is fogok
 yesterday already meet-past-1sg with the boys who-with today also will-1sg
[VP *találkozni*]
meet-inf
 ‘Yesterday I already met the boys I will (meet) today, too.’
- (d) *Holnap is fogok *találkozni* azokkal a fiúkkal, akikkel tegnap
 tomorrow also will-1sg meet-inf with those boys who-with yesterday
[VP *találkoztam*]
meet-past-1sg
 ‘Tomorrow I will meet the boys I did (meet) yesterday, too.’

Somewhat differently, though, there are cases of subordinating constructions where tense-marking differences are also licit in ellipsis:

- (4) Holnap is ugyanazt *játsszuk* amit tegnap [VP *játszottunk*]
 tomorrow also the same play-pres-1pl what yesterday play-past-1pl
 ‘Tomorrow we will play the same as we did yesterday.’

Not willing to get entangled into even more details, though, we will confine ourselves to coordinating constructions, as has been put forth in the introduction, and abstract away from these cases.

The basic generalizations that offer themselves are summarized in (5):

- (5) LICENSING CONDITIONS FOR VP-ELLIPSIS IN HUNGARIAN
- Agr_x licenses the deletion of Agr_y
 - T_x licenses the deletion of T₀ (= infinitive)
 - T_x does not license the deletion of T_y if $x \neq y$
 - T₀ does not license the deletion of T_x unless $x = 0$

These hold as well as the most fundamental requirement: that the V-stem must be recoverable, too, e.g., as licensed by an identical antecedent stem.

2. Lasnik on English VP-ellipsis

Lasnik (1995), explicating on the difference between lexical and syntactic affixation, examines the behavior of English verb forms under VP-ellipsis. He notes, following Quirk et al. (1972) and Sag (1976), that a bare V-form can

be elided after a modal when anteceded by a tensed V-form (6a, b), or even by a participle form (6c, d):

- (6) (a) John *slept*, and Mary will [_{VP} ~~sleep~~], too.
 (b) John *sleeps*, and Mary should [_{VP} ~~sleep~~], too.
 (c) ?John *was sleeping*, and Mary will [_{VP} ~~sleep~~], too.
 (d) John *has slept*, and Mary will [_{VP} ~~sleep~~], too.

It might then be proposed, writes Lasnik, that sloppy identity is at work here, but not only for ϕ -features, where it is quite common, but also for tense/aspect marking. However, there are data which suggest otherwise:

- (7) (a) *John *was/is here* and Mary will [_{VP} ~~be here~~], too.
 (b) John will *be here*, and Mary will [_{VP} ~~be here~~], too.
 (c) *John *has left*, but Mary shouldn't [_{VP} ~~have left~~].
 (d) ?John should *have left*, but Mary shouldn't [_{VP} ~~have left~~].

As (7b) shows, there is no general problem with eliding *be*, yet in (7a), contrary to what we have seen in (6), the tensed form of *be* cannot properly antecede the bare form. Basically the same contrast is observable with respect to *have*: (7c, d). In these cases then it is strict, and not just sloppy, identity of verb forms that is the prerequisite of ellipsis. Note, further, that this identity condition holds for abstract, featural forms, rather than surface phonological forms:

- (8) *The men *have left*, but the women shouldn't [_{VP} ~~have left~~].

In (8), the antecedent and the elided verb share the same surface form, but they differ as to their abstract feature content: the antecedent in the first clause is a tensed (though unmarked) form, whereas the elided one in the second clause is a bare infinitive. The fact that the ellipsis fails suggests that surface identity is not sufficient.

Lasnik thus draws a distinction between main verbs on one side, and *have* and *be*¹ on the other, stating that with the former sloppy identity is the relevant

¹ Observe that *be*, whether a main verb, or an aspectual auxiliary, behaves the same way throughout.

criterion, while the latter require full formal identity. Following Sag (1976), however, Lasnik shows that the strict identity requirement on ellipsis can be maintained for all cases, on the assumption that there is a difference between main verbs and auxiliaries as to the phase, in their morphological formation process, where ellipsis is carried out in the derivation. Auxiliaries are taken to be lexically affixed, i.e., they bear all inflection already at the point when they enter the syntactic derivation. Main verbs, on the other hand, are inserted unaffixed, and attach to their inflectional associates only in the course of syntactic derivation. (*Be* is an auxiliary even when it is a main verb at the same time: it undergoes the same processes (raising, inversion, etc.) that characterize the group of auxiliaries in English.) The ellipsis facts seen above follow now if the deletion of the VP can take place before the main V attaches to its inflectional ending: At that point the antecedent is still a bare V-form, hence strict identity permits the parallel deletion of another bare V-form. This option, however, is not available for auxiliaries, since they never occur in syntax without their (lexically fixed) inflection—so the only ellipsis possibility arises when the two V-forms in question are fully strictly identical, including the inflectional part. Some illustration is given in (9):

- (9) (a) John [_{INFL} -s] [_{VP} sleep] and Mary [_{INFL} will] [_{VP} ~~sleep~~], too.
 (b) *John [_{INFL} is] [_{VP} t here] and Mary [_{INFL} will] [_{VP} ~~be here~~], too.

Lasnik's account also predicts that a bare form can only be deleted in an analytic construction, i.e., one in which Infl is filled by an independent word, otherwise the deletion of the bare V, that is, the stem, would strand any affixal Infl, which would result in an ill-formed structure.

3. Hungarian: Strict or sloppy identity?

We now return to the Hungarian data, to try to determine what conditions govern the deletion of V-forms in VP-ellipsis. At first sight, it appears that we are very close to settling for a sloppy identity analysis. Agr-features can widely differ between the antecedent and the elided V, whereas tense-marking is not flexible in this respect—more or less the state of affairs expected for sloppy identity, which normally only affects ϕ -features, i.e., Agr. There are two points, though, the consideration of which may press us to attempt to reach behind sloppy identity. The first point is the conceptual issue: Obviously,

an account based on strict identity is more constrained, hence preferable on general metatheoretical grounds. The second point has to do with the tense-marking effects, in particular with the fact that an elided infinitive is rendered legitimate by a tensed antecedent, but not the other way round. A sloppy identity analysis will say absolutely nothing about this, since, as alluded to above, its main motivation would consist in the fact that it does not involve tense-marking. Now, if a correct analysis can be suggested which could cover all aspects of the data (Agr-sloppiness as well as tense-strictness) making use of the notion of strict identity, it would be superior to the (otherwise almost self-suggesting) sloppiness account both conceptually and empirically.

Let us first test Lasnik's theory on the data. Recall that the key to that analysis was the availability of some V-form (most importantly the base) in morphosyntactic isolation at a point in the derivation where the ellipsis/deletion could be carried out. With lexically affixed forms only the whole V-form could serve as the strictly identical antecedent, while with syntactically assembled forms if the deletion could precede the affixation, some stem form could be accessed.

Even before trying to establish whether affixation is lexical or syntactic, Hungarian poses an immediate problem. In this language no final V-forms are base forms. The infinitive bears an affix (*-ni*, as in *fut-ni* 'run-inf: to run'), and the most unmarked form, 3rd person singular in present tense, albeit phonologically indistinguishable from the base, is nevertheless morphologically complex: it bears at least an agreement suffix, even if it is phonologically null in the vast majority of cases.² As you may remember from section 2, what counts is the abstract, morphological or featural make-up, not the surface form, so the 3SG forms will not be identical to the base at the relevant level of representation. Thus we have a problem: Even if assuming syntactic affixation, there is no hope to ever find an appropriate candidate for ellipsis in a sloppy agreement identity case, since the largest chunk that could be deleted under strict identity is the base, but if it gets deleted, at least the agreement affix will be left behind stranded, in violation of generally accepted principles.

To overcome this problem, the minimal solution seems to be to propose that at the level of PF a morphological **stray erasure** (similar to its phonological counterpart generally assumed) is operative, deleting stranded affixes, as long as the (equally usual) condition of recoverability is observed.

² There are cases, though, where this 3sg agreement suffix is overt, e.g., the *-ik* ending in a particular inflectional V-type: *esz-ik* 'eat-s', or the *-On* ending in some archaic forms: *lesz-en* 'will_be-s'.

With this auxiliary hypothesis, we can now examine how Lasnik's insights could be applied to our data. In Hungarian, as opposed to English, the obvious distinction to be drawn is between different affix-classes, rather than stem-classes. Unlike English, Hungarian verbs display uniform behavior, but it seems that tense and agreement affixes belong to distinct levels of application. The core idea is to take tense affixes to be lexically attached to stems, hence the lack of variability between the elided form and its antecedent: the base forms are uniformly unavailable for syntactic operations already. On the other hand, agreement affixes are syntactically attached, hence the [V+T] stems, still separate from Agr-features, can be targeted by ellipsis under strict identity — the only thing that needs to be looked after by some other mechanism is the elimination of the Agr-affixes rendered hostless by the deletion.

As regards the case of infinitives, they must be looked upon as forms unmarked for tense, so when they undergo ellipsis, deletion can apply regardless to the tense-marking of the antecedent, as the infinitive is a mere base (disregarding Agr-features for the moment, since they are irrelevant here), so it is identical to a subpart of the antecedent [V+T] form (or fully identical with an antecedent infinitive). But when the intended antecedent is an infinitive, it cannot license the ellipsis of a tensed, lexically formed [V+T], since the antecedent contains less information than the target, thus recoverability would be violated.³

While the solution just outlined certainly has some appeal, it can still be proved to be inferior to some other analysis. For one thing, we needed a stipulation about lexically and syntactically attached affixes, plus an auxiliary hypothesis concerning the necessarily stranded Agr-affixes. For another, in the next section further data will be introduced to show that the Lasnik-style account faces even more serious empirical problems. Therefore I now turn to the presentation of a new analysis, which is hoped to overcome the difficulties mentioned.

³ It may even be argued that the morphological template of the infinitive lacks T altogether, i.e., it is merely V(+Agr), so it is not even structurally equivalent to a tensed form, hence its incapability of serving as a legitimate antecedent.

4. The joys of late insertion

4.1. Object agreement: A challenge

Halle and Marantz (1993) put forth a theory of the syntax–morphology interface in which syntax is assumed to operate on features and feature bundles rather than full lexical items with phonological shapes. The actual forms to be interpreted phonologically are inserted at a later point in the derivation, more notably after s-structure (or spellout), on the branch of the derivation targeting PF. They propose that a separate representational level exists for morphology and lexical insertion (they call it Morphological Structure (MS)), where a list of lexical forms (the Vocabulary) is accessed, to fill out the morphologically determined terminal nodes of phrase structure trees. In this matching procedure several possible candidates compete for the insertion points, and the one that most specifically matches the feature bundle at the terminal node gets inserted.

Before this so-called **late insertion** takes effect, certain morphological manipulation is possible (and in fact necessary) on the terminal nodes of the syntactic structure, such as the merging of (the feature content of) adjacent nodes, or the fission of a node into multiple insertion points, etc. These are needed to cater for well-known mismatches between morphemic structure and morphophonological structure.

Clearly, under this view of the organization of grammar, what are traditionally referred to as deletion processes (including ellipsis) are cases of **non-insertion** at MS, i.e., at the place of the “elided” part the terminal feature complexes will not get associated with phonologically interpretable content (and the whole phenomenon will not have any consequence for LF, since whatever information pertains to LF is sent there in the form of feature bundles only, on the other branch of syntactic derivation).

Perhaps the most compelling evidence in favour of such an approach to the problems of ellipsis in Hungarian comes from object agreement data. In this language verbs show a certain form of agreement with the object, as well as the subject. The content of subject agreement is ϕ -features, as has been shown above, but object agreement is different in nature. Traditionally it is referred to as “definiteness” agreement: if the object nominal is definite, it triggers a special Agr-paradigm called the “objective conjugation”, i.e., there is no(t always a) separately identifiable affix representing object agreement, it is rather a choice between different subject agreement paradigms. Recently, Bartos (1997) argued that object agreement is a categorial matter: objective conjugation obtains if and only if the object is a DP by category, and when it

is of a lesser type (NP or NumP), or when there is no object at all, subjective conjugation is found on V. But whatever the true nature of this object agreement, we find that under ellipsis it behaves basically like subject agreement, i.e., it can be freely varying between the antecedent and the deleted V, cf. (10):

- (10) (a) Péter az összes kutyát látta de mi csak hármat
 Peter the all dog-acc saw-3sg-obj_conj but we only three-acc
 [VP látunk].
 saw-1pl-subj_conj
 'Peter saw all of the dogs, but we only saw three (of them).'
- (b) ?Péter csak három kutyát látott de mi mindet
 Peter only three dog-acc saw-3sg-subj_conj but we all-acc
 [VP láttuk].
 saw-1pl-obj_conj
 'Peter only saw three (of the) dogs, but we saw all (of them).'

On one hand, this is no surprise given that object agreement is not strictly affixal, but featural, influencing the paradigm choice for subject agreement (Agr). On the other hand, there are types of cases where an affix surfaces in objective conjugation which can most readily be analysed as the exponent of object agreement (see (11) for an example).

- (11) (a) vár-∅ vár-ja-∅
 wait-3sg wait-obj_agr-3sg
- (b) vár-tok vár-já-tok
 wait-2pl wait-obj_agr-2pl
- (c) vár-t-∅ vár-t-a-∅
 wait-past-3sg wait-past-obj_agr-3sg
- (d) vár-t-atok vár-t-a-atok (→vártatok)
 wait-past-2pl wait-past-obj_agr-2pl

But whatever the correct segmentation or analysis, the syntactic background raises a serious problem here. As often assumed to hold universally, and specifically argued to hold in Hungarian by Brody (1995), the syntactic hierarchy of inflectional projections is Agr_SP > TP > Agr_OP. Also, there are reasons to maintain Baker's (1985) Mirror Principle, which states that morphological derivations (hence the order of affixes) reflect syntactic derivations (i.e., on the assumption of head movement creating the relevant morpheme sequences, the projectional hierarchy). In particular, then, what we expect the situation to

be is the morphemic order $V + \text{Agr}_O + T + \text{Agr}_S$, as the verb raising successively to the syntactic head positions picks up the affixal heads in this order. We now face two difficulties. First, whether object agreement is affixal or just a feature of subject agreement paradigm choice, it is in the “wrong” place morphologically: as an affix it should precede T , or as a feature, it should be incapable of merging with Agr_S , skipping T . Secondly, if Agr_O is stuck between V and T , and we assume in Lasnik’s (1995) mood that $[V+T]$ is a lexically fixed unit, then Agr_O , being syntactically closer to the stem than T , should also be lexically attached (or else checking between V and the inflectional heads could not proceed in the required way), but then Agr_O should pattern with tense in its behavior under ellipsis—which is not the case, as has been shown in (9).

A satisfactory discussion of the first problem largely exceeds the scope of this paper, and (hopefully) we can set up our analysis of ellipsis without having a final solution for it at hand. Bartos (2000) is meant to shed some light on the issue, but now we leave this question open, noting that in a more favorable case solving the first problem might find the second one gone entirely. Nevertheless, even if we ignore this empirical problem for Lasnik’s (1995) analysis, his theory faces further difficulties datawise (see below), and has certain conceptual drawbacks, as has been pointed out at the end of Section 3.

4.2. Licensing factors: Parallelity and matching

Let us consider now, what theoretical options we are left with. We have seen that Lasnik’s (1995) account is possibly viable but can (and therefore should) be improved upon. The question is which way we go with the improvement process. The minimalist framework is a lexicalist model at heart. This means that all word forms handled by syntax are atomic morphological units, whose inner structure is inaccessible for syntactic and post-syntactic processes. While it is quite possible that our Hungarian data could be covered by a lexicalist theory, Lasnik’s results presuppose a model in which syntactic affixation is an option, that is, a non-fully-lexicalist model, by definition. Since our aim is to maintain the validity of Lasnik’s insights, we must look for a solution outside the domain of strict lexicalism. Halle–Marantz’s (1993) model appears to be a suitable one, insofar as it provides ways to capture Lasnik’s lexical vs. syntactic attachment distinction, while at the same time it paves the way for an account of the Hungarian data.

To see how things can be made to work, consider a particular case:

- (12) Péter alszik és én is [_{VP} alszom].
 Peter sleep-3sg and I also sleep-1sg
 'Peter is asleep, and so am I.'

This is the typical borderline case, where there is no tense difference, only agreement difference. At the point where the derivation feeds the structure into MS, the verb complexes have already been assembled, i.e., assuming a minimalist-type derivation, V has already picked up T and Agr_S successively. So categorially, what we loosely call V is [_{Agr_S} [_T V T] Agr_S], in fact, occupying Agr_S⁰.⁴ (Consequently, though essentially irrelevantly, what we loosely refer to as VP-ellipsis is usually Agr_SP-ellipsis, but it should not cause pains, Agr_SP being an inflectionally extended VP, in the sense of Grimshaw (1991; 1997).⁵) The actual stem and affixes are still represented by feature bundles. When it comes to Vocabulary insertion at MS, that is, the phonological forms get associated with the featural terminal nodes, insertion proceeds item by item. Each time there is a consideration of whether non-insertion (i.e., ellipsis) is an option. Surely, the decision hinges on whether the omission is licensed or not. Arguably, whenever there is a parallel construction in the structure with identical parts, the possibility of omission under recoverability arises. For the V-stem and the tense affix this is necessary and sufficient. For Agr, it would be sufficient, but obviously isn't necessary. The reason for this is probably the fact that Agr is supported by (dependent on) feature matching—the very nature of agreement. In other words, there must be some element present in the structure that V agrees with, putting the fact on display in the form of the Agr-affix.

It must thus be the case that apart from the parallel identity licensing of ellipsis, another factor may be involved, that of licensing by featural agreement. V and T may be elided if, and because, there is an identical element occupying a structurally parallel slot. Agr may be elided either in the same case, or if, and because, it is identifiable, hence recoverable, by the element agreeing with its host, the verb. In essence, then, when a V+T+Agr form undergoes ellipsis, the V+T part does this under a different condition than the Agr part. The

⁴ In Halle – Marantz's (1993) model, the V+Infl complexes must be formed before the derivational branching to LF and PF, i.e., 'overtly', so if it is done via verb raising, it cannot be left to after the branching, unlike in Chomsky's (1995) minimalism. Apparent counterexamples, such as the "affix-hopping" nature of English inflexion, are handled by morphological merger under adjacency. In Hungarian, however, V always raises to at least Agr_S overtly (Szabolcsi 1996), so the problem will not arise for us here.

⁵ Cf. also the discussion of the category of ACD-targets in Hornstein (1995, 87ff).

omissibility of Agr, then, is much wider in scope than that of V or T, since Agr itself is about licensing by matching, i.e., whenever it appears, some matching must have occurred, creating the fundamental possibility of Agr-ellipsis. Yet, Agr-ellipsis may never occur in itself, even though the licensing condition is automatically met, except when the other participant of the matching (the item that V agrees with) is elided, in which case the omission of the Agr-suffix would eliminate the only trace of the fact of matching.⁶

As (13) shows, pure Agr-ellipsis is illicit, although the agreeing item (*én* 'I') is present as a potential licenser, so the omission would be recoverable:

- (13) *Tegnap én vit-t-[em] le a szemetet.
 Yesterday I take-past-1sg down the garbage
 'Yesterday I took the garbage down.'

To account for this fact, we must resort to some principle of word integrity, something like an **all-or-nothing** condition:

- (14) ALL-OR-NOTHING CONDITION

No proper part of a word may be affected by ellipsis.

Note, furthermore, that the **strict identity condition** on ellipsis is observed in a special sense by the omission of Agr, as well. Only, it need not be strictly identical to a parallel antecedent, but to the ϕ -feature bundle of the anchor of agreement. Recall that identity is assumed to operate on an abstract, featural level, and that syntactically Agr-affixes are pure ϕ -feature bundles.

As regards object agreement, which was seen to behave like subject agreement under ellipsis, it must be licensed by local matching whenever the elided VP has a different specification for Agr_O from that of the parallel remaining VP. This is easily achieved, on grounds that object agreement is governed by some property (presumably the projectional category) of the object. So if the object (or the lack of any object) identifies the Agr_O morpheme of the elided VP, the structure is recoverable, hence licit, not in need of relying on identity with the parallel structure.

⁶ In Hungarian, subject personal pronouns are usually pro-dropped, exactly because the Agr-ending on V may identify this pro. They must be overt, however, when they are focused or contrastively topicalized, and VP-ellipsis is strongly linked to focusing, so it is no surprise that the prime examples for the overt subject identifying the missing Agr-suffix are our VP-ellipsis cases.

Next consider the ellipsis of an infinitival form as licensed by a parallel tensed form:

- (15) Péter aludt és én is fogok [~~aludni~~].
 Peter sleep-past-3sg and I also will sleep-inf
 'Peter has slept, and I will (sleep), too.'

The morphological make-up of the elided form is $V + Af_{inf}$, that of the licenser is $V + T_{past} + Agr_{3sg}$ (the Agr-affix is phonologically zero). The key question regards the status of the infinitive-marking. If it is a realization of the syntactic category T^0 , it poses a problem, because then we have two different tenses (T_{past} vs. T_{zero}), so the licensing condition we assume is not directly met. Two options offer themselves. Either the analysis of the form must be revised, or a different licensing mechanism is at work. Let us explore these in some detail.

Reconsidering the make-up of infinitives, we may propose that the suffix *-ni* has no morphemic status. In other words, the morphologically complex form $V + ni$ realizes a syntactically simplex unit: V unspecified for the feature [tense]. In Halle–Marantz's (1993) **Distributed Morphology** (DM) model there exists a morphological operation responsible for such cases: **fission**, whereby a feature bundle breaks up into two, so that it will yield two nodes for Vocabulary insertion.⁷ In this particular case, then, ellipsis is licensed for a V -form unassociated with T —in essence, the V -stem of the parallel clause serves as its identifier, regardless whether this antecedent is or is not associated with T . In the infinitival clause T is either missing altogether, or more probably, it is phonologically empty, i.e., nothing is inserted there at MS anyway, so elliptical non-insertion considerations are entirely irrelevant.

The second alternative for the problem illustrated by (15) would be to refer to another licensing mechanism in this case. In particular, we may claim that when an infinitival form is elided, somewhat similarly to what we have seen for Agr-affixes, the [$-tense$] affix is licensed for ellipsis separately from its stem. The V -stem is licensed in the usual way of parallelity, whereas the affix is licensed by another syntactic relation: selectional dependence. Recoverability is satisfied by way of the selectional relations in the structure, since the infinitival forms in the relevant examples are uniformly selected by a superordinate V or Aux, thus the elided V -form can be unambiguously identified as an infinitive.

⁷ The occurrence of fission is probably motivated here by the tendency of avoiding homonymy: Absolutely bare V -forms do not surface in Hungarian; the least marked form is the 3sg present form, with phonologically null tense and agreement marking. The infinitive is overtly affixed exactly to avoid surface identity with this inflected form.

Whichever of the two solutions proves to be correct eventually, the cases of this type are accounted for without affecting the analyses of the other types. The reverse cases (infinitives unsuccessfully trying to antecede ellipses of tensed forms) fall out in either scenario. If infinitives are syntactic simplexes, then the elided tense-marking in the other clause will be irrecoverable—it is neither properly anteceded, nor licensed by any agreement or selectional relation. If, as in the second option, infinitives are complex, the situation will be exactly the same for the tensed ellipsis targets.

4.3. Reinventing the wheel: Reproducing Lasnik's results

What remains for us to do is to show how our theory can replicate Lasnik's (1994) results. Recall that the crucial point in his analysis of English inflection was the distinction between lexical affixation of Aux's (*have*, *be*), and syntactic affixation of main verbs. This way, main V stems were accessible ellipsis antecedents under strict identity at a point of syntax when V and Infl are still separated, hence the availability of 'sloppy identity'. With auxiliaries, however, only full V+Infl identity suffices, since these forms are complex already at the point when they enter syntax.

In our DM-based model this distinction cannot be directly implemented: All forms must be assembled by the time of Vocabulary (non-)insertion, where ellipsis is determined, since there is no place for further syntactic derivation beyond that. Nevertheless, the phenomenon can be captured, for instance in the following way. Assume that the key difference between *have*, *be* and main verbs is that the former must raise to T, while the latter remain in a lower position (V, or at best Agr_o).⁸ Auxiliaries thus reach the entry point to MS as adjoined to their inflection, and the two nodes fuse at MS. (Recall that the difference between merger and fusion is that in the former the feature bundles are kept separate, and two nodes remain for insertion, forming one word-domain; and in the case of fusion the two feature bundles are unified, and only one insertion node remains available.) So when ellipsis possibilities are determined, auxiliaries are indivisible featural units, fully specified for tense,

⁸ I will not discuss the background of this difference in behavior; for some discussion see Lasnik (1995), Chomsky (1995)—the latter attributes this difference to their semantics, the former claims that the distinction is lexical: all inflected forms of auxiliaries are stored separately there, because of their highly idiosyncratic nature, while main verbs are stored as roots, and their (more or less) regular tense suffixes are stored as separate entities, too.

therefore only fully identical forms (matching in stem as well as in inflection) are accessible for recoverability, the crucial prerequisite of ellipsis.

On the other hand, main verbs do not raise to T; they enter MS as separate nodes, therefore the stems are available antecedents for the same stem forms in parallel structures for omission. V and T will only get morphologically associated by the MS-process of **merger**, but the feature bundles remain distinct, and insertion of V and T are two distinct operations. (But of course they go together insofar as word integrity must be respected, cf. the **All-Or-Nothing Condition** in (14) above.) Also, if there is any material intervening between T and V, merger cannot take place, so T and V are lexicalized absolutely independently of each other—in such cases T can only be realized by a modal, or a supportive *do*, carrying inflection.

Thus Lasnik's results can be recaptured in this DM-based model, too, without any extra conditions or stipulations that would otherwise be unnecessary.

5. Conclusion

I have treated VP-ellipsis phenomena from the viewpoint of the behaviour of verbal inflection under ellipsis. I started out from Lasnik's (1995) analysis of English data, and provided a range of relevant Hungarian data, too, to show that Lasnik's account does not simply carry over to them. Therefore I proposed and explored another analysis, embedded in a model of syntax constructed on the basis of Distributed Morphology, as put forth by Halle-Marantz (1993). This account has a clear conceptual advantage over Lasnik's one, as it does not need to appeal to the notion of stray erasure in order to handle the Hungarian data properly. Furthermore, in certain cases (such as the problem of object agreement, and the word integrity effects) it is directly capable of covering the data, whereas Lasnik's theory either fails completely (depending on the availability of solutions to the object agreement problem), or needs extra hypotheses to go through. To give further credibility to my model, in the following Appendix I will sketch how it can be employed in accounting for other typical ellipsis effects, now routinely treated in LF-reconstruction terms.

Appendix

Although the present paper is concerned with the behaviour of verbal inflection under ellipsis, and not any other aspect, such as pronoun interpretation, I find it important to say at least a few sentences about a possible direction of developing the ideas pursued above, so as to find an account for the classic domain of strict vs. sloppy identity: pronoun interpretation. This is one of the motivating areas of the recent LF-reconstruction approaches to ellipsis phenomena (see e.g., Fiengo–May 1994; Hornstein 1995; Kitagawa 1991). In what follows, I will give a brief illustration of how the relevant data might be handled in my DM-based, essentially deletional (rather than reconstructional) model. I will rely on Fiengo–May’s (1994, 129ff) material for my presentation.

Take an example to illustrate what problems we face:

(A1) Max saw his mother, and Oscar did [_{VP} see ~~his mother~~], too.

This is probably the best-known example of the strict vs. sloppy reading issue: The elided VP can be understood either as “see Max’s mother” (i.e., Oscar saw Max’s mother)—the strict reading, or as “see Oscar’s mother”—the sloppy reading. The reconstruction theories assume that throughout the syntactic derivation the VP of the second clause is empty, and is filled in (“reconstructed”) only at LF, where the referential identity of the pronouns gets fixed, too. Put in elementary terms, what happens is that the pronoun *his* is literally reconstructed on the basis of the first clause, and either before, or after this copying, indexes get assigned. If index assignment occurs first, the pronoun in the second clause necessarily inherits the index of the pronoun in the first. If the order of the operations is reversed, the reconstructed pronoun is free to receive any index (so that it either refers to Max, or to Oscar, or to a third party).

As opposed to this conception, my model has to assume that the different readings of (A1) correspond to different sentences from the very beginning: The features (among them: the index features) are different in the differently interpreted cases throughout the derivation, and the question for me is not why there are these many construals, but when ellipsis is licit.

With respect to (A1), the following readings may arise:

(A2) Max saw his_M mother, and Oscar did [_{VP} see his_M mother]. – anaphoric, strict

(A3) Max saw his_M mother, and Oscar did [_{VP} see his_O mother]. – anaphoric, sloppy

(A4) Max saw his_X mother, and Oscar did [_{VP} see his_X mother]. – $X \neq \{M \vee O\}$; deictic, strict

(A5) Max saw his_X mother, and Oscar did [_{VP} see his_Y mother]. – $X \neq Y$, deictic, sloppy

In fact, (A5) is out, even though in a non-elided case (*'Max saw his mother, and Oscar saw his mother.'*) it is clearly possible to use both instances of *his* deictically referring to two different individuals. The correctness of an analysis turns on exactly this case. While it is not trivial for a reconstructional theory to yield this result, either (see Fiengo–May 1994, Ch.4., for detailed discussion), I am now concerned with how our model can cater for it. Basically, in this model all four sentences are different for the computational system, with respect to the index feature of the pronouns. Both clauses are present from the starting point of the derivation through LF. Ellipsis occurs on the PF-branch, as non-insertion of phonological material. I must explain, why VP-ellipsis is allowed in (A2–4), but not in (A5). In (A2), the strict anaphoric construal of *his*, the omission of the pronoun⁹ is licensed by the parallelly existing featurally identical form. In (A3), the sloppy anaphoric case, the pronoun is recoverable by its anaphoric relation to 'Oscar', the most local antecedent. In (A4), where both pronouns refer deictically to a third individual, parallelity licensing is once again invoked. But in (A5), the doubly deictic, unavailable case, *his* in the second clause cannot be fully recovered if deleted, since it is neither identical to its parallel in the first clause, nor anaphorically linked to 'Oscar'; therefore, no wonder, ellipsis is rendered illicit—the correct result.

Let us see now a more complicated case cited by Fiengo–May (*ibid.*):

(A6) Max said he saw his mother, and Oscar did [_{VP} ~~say he saw his mother~~], too.

Disregarding here the deictic cases, there are four possibilities, once again:

(A7) ... Oscar did [say he_M saw his_M mother]. – strict ATB

(A8) ... Oscar did [say he_O saw his_O mother]. – sloppy ATB

(A9) ... Oscar did [say he_O saw his_M mother]. – sloppy + strict

(A10) ... Oscar did [say he_M saw his_O mother]. – strict + sloppy

Once more, the last option (A10) is sharply out (though (A9) is quite marginal, too). That is, we can have strict or sloppy identity across the board (ATB), but cannot switch from strict to sloppy on the fly, albeit the reverse is possible.

⁹ Obviously, the elision of the rest is straightforwardly licensed by the parallel clause, so I do not discuss it any further.

The ellipsis in (A7) is licensed by the parallel clause, the one in (A8) is licensed anaphorically, for both pronouns. (A9) shows that the first elided pronoun can be recovered as anaphoric to ‘Oscar’ (the most local potential antecedent), while the second is identified under parallelity. But in (A10), only the first one can be elided legitimately (under parallel licensing)—the second pronoun cannot rely on the parallel clause, but nor can it be properly linked to ‘Oscar’ as an antecedent, exactly because there is an intervening potential antecedent: *he*, which, however, is linked to ‘Max’, by parallelity, hence if the second pronoun was recovered anaphorically, the sentence would fall in with (A7), which it should not. Thus, once again, the unavailable option is ruled out by our model correctly.

At this point I quit the discussion of the pronoun interpretation issue, since it is not in the focus of this paper. I hope that this brief illustration has shown that our model has the potential to cope with this kind of data, relying on a deletional, rather than reconstructional, treatment of ellipsis, advocated for the DM-based analysis of verb form omissions.

References

- Ackema, P. 1995. Syntax below zero. Doctoral dissertation, Utrecht University.
- Baker, M. 1985. The Mirror Principle and morphosyntactic explanation. *Linguistic Inquiry* 16: 373–415.
- Bartos, H. 1997. On “subjective” and “objective” agreement in Hungarian. *Acta Linguistica Hungarica* 44: 363–84.
- Bartos, H. 2000. Az inflexiós jelenségek szintaktikai háttere [The syntactic aspects of inflectional phenomena]. In: Kiefer, F. (ed.): *Strukturális magyar nyelvtan 3. Morfológia. [A structural grammar of Hungarian 3. Morphology]*, 653–762. Akadémiai Kiadó, Budapest.
- Bródy, M. 1995. Focus and Checking Theory. In: Kenesei, I. (ed.): *Approaches to Hungarian. Volume 5: Levels and structures*, 29–41. JATE, Szeged.
- Chomsky, N. 1995. *The Minimalist Program*. The MIT Press, Cambridge MA.
- Fiengo, R. – May, R. 1994. *Indices and identity*. The MIT Press, Cambridge MA.
- Grimshaw, J. 1991. *Extended projection*. Ms., Brandeis University.
- Grimshaw, J. 1997. Projection, heads, and optimality. In: *Linguistic Inquiry* 28: 373–422.
- Halle, M. – Marantz, A. 1993. Distributed morphology and the pieces of inflection. In: Hale, K. – Keyser, S.J. (eds): *The view from Building 20. Essays in linguistics in honor of Sylvain Bromberger*, 111–76. The MIT Press, Cambridge MA.
- Hornstein, N. 1995. *Logical Form*. Blackwell, Oxford & Cambridge MA.

- Kitagawa, Y. 1991. Copying identity. *Natural Language and Linguistic Theory* 9: 497-536.
- Lasnik, H. 1995. Verbal morphology *Syntactic structures* meets the minimalist program. In: Campos, H. – Kempchinsky, P. (eds): *Evolution and revolution in linguistic theory: essays in honor of Carlos Otero*, 251-75. Georgetown University Press, Washington D.C.
- Quirk, R. – Greenbaum, S. – Leech, G. – Svartvik, J. 1972. *A grammar of contemporary English*. Longman, London.
- Sag, I. A. 1976. *Deletion and Logical Form*. Doctoral dissertation, MIT, Cambridge MA.
- Szabolcsi, A. 1996. *Verb and prefix movement in Hungarian*. Ms., UCLA.

Address of the author: Huba Bartos
Research Institute for Linguistics, HAS
and
Theoretical Linguistics Doctoral School, ELTE
H-1399 Budapest, P.O.B. 701/518.
bartos@nytud.hu

COMPLEMENT SELECTION IN MORPHOLOGY AND SYNTAX*

ANTONIETTA BISETTO AND SERGIO SCALISE

Abstract

In this paper we discuss a recent proposal according to which: (1) derivational suffixes have a syntactic-like bar representation and (2) word formation processes can be represented in terms of an X-bar syntax. While we agree on the attribution of an argument structure to suffixes, we cast some doubt on the claim that such word-constituents are Complements of their (suffixal) selecting heads. We maintain that suffixes do not project as lexical heads do and that a configurational X-bar structure in word formation is useful only to the extent that it represents the semantics (i.e., the LCS) of the words selected by the suffix.

1. Short history of lexical and affixal representations

The representation of affixal heads has often been modelled on the representation of lexical heads. This has been true for many reasons, among which: (1) the representation of lexical heads has always been part of the concern of syntax, i.e., since affixal heads give rise to words (= lexical elements), which are relevant to syntax, affixes require the same representation as lexical heads; (2) the parallelism between the representations of the two types of heads seemed to be necessary in order to guarantee the reciprocal matching of the two items: the base and the affix. In the recent past, two main proposals have been put forth in order to account for the representation of both lexical and affixal heads, as well as the matching possibilities between the two, namely, representations in terms of: (a) subcategorization frames and (b) theta grids.

* This research has been carried on with the support of the Italian CNR. We would like to thank Donna Miller (University of Bologna) for checking the English version. This paper is the result of collaboration between the two of us. However, the final responsibility for sections 1-3 is of Sergio Scalise and for sections 4-7 of Antonietta Bisetto.

1.1. Subcategorization frames

Subcategorization frames were introduced by Chomsky (1965) according to whom “[...] Strict subcategorization specifies the categorial nature of the complement of a lexical head”. Strict subcategorization was then adopted in derivational morphology explicitly, for example by Lieber (1980; 1983). However, Generative Morphology, globally considered, generally speaking (cf. Aronoff 1976; Scalise 1984) has opted to represent subcategorization frames as in (1), even though the element an affixal head subcategorises for (V in the example below) cannot be considered a complement.

- (1) *lexical* *affixal*
drive [___ NP] -er [V ___]

Strict subcategorization has been part of the representation in use since *Aspects* (Chomsky 1965) and was coupled with the features that that model used to represent the syntactico-semantic characteristics of an item, as illustrated in (2):

- (2) *lexical* *affixal*
vino] N [−abstr] -aiο] N [+um] → vino
'wine' 'wine seller'
gioia] N [+abstr] -aiο] N [+um] → *gioiaio
'joy' *'joy seller'

where it can be seen that the suffix *-aiο* selects Nouns characterised by the feature [−abstract].

1.2. Theta-grid representation

In the late eighties, Roeper (1987) proposed to extend to affixal elements the theta-grid representations that had been proposed for lexical items after Gruber (1967) and Jackendoff (1972):

- (3) *lexical* *affixal*
[V (Ag,Th)] -able (Ag,Th)

The representation in (3) shows that the suffix, which is represented as having a theta grid “Agent Theme”, selects verbs with a theta grid specified for the same theta roles and forms Adjectives with the same theta grid.

2. Comments and criticisms

The two kinds of representations above, however, proved inadequate to account for the selection a suffix makes on the words which it can attach to. Neither of them, in fact, was able to solve the problem of adequacy in word formation.

Affixal subcategorization does not work well because, as shown by Di Sciullo (1992; 1993; 1996), it is based on a head-to-head relation which does not allow the suffix to access certain information characterising the base word. Such information can be captured only representing the word by means of a projection structure. This can be seen with the Italian suffix *-tore* ‘-er’, a suffix which selects Verbs in the same way as its English counterpart.¹ *-tore* selects verbs which can be either transitive or intransitive. The transitive/intransitive specification, however, does not suffice for the correct attachment of the suffix because such information does not differentiate between ergative and unergative intransitives. Even if subcategorization were to be refined with features such as [+ergative] and [-ergative], as Di Sciullo argues, it would still be insufficient, because there is no way of exactly restricting the attachment of the suffix.

As the examples in (4) show, *-tore* cannot be attached to any kind of transitive or intransitive verbs:

- | | | | |
|-----|---------------------------|--|-----------------------|
| (4) | <i>transitive</i> | | <i>intransitive</i> |
| (a) | trasportatore ‘carrier’ | | corridore ‘runner’ |
| | consumatore ‘consumer’ | | fruitore ‘user’ |
| (b) | *preoccupatore ‘worrier’ | | *arrivatore ‘arriver’ |
| | *disgustatore ‘disguster’ | | *dormitore ‘sleeper’ |

The properties of verbs not allowing the formation of the *-tore* words in (4b) cannot be grasped if the affix is allowed to ‘see’ only the category information of the head (V); the representation of the base constituent in purely categorial terms is indeed inadequate to prevent a free attachment of the suffix. *-tore* actually needs verbs whose subjects are characterised by semantic features like ‘agentivity’ and/or ‘intentionality’ (as we will see in detail below), and features such as these are available only if a more complex representation is supplied.

Summing up, representations with subcategorization frames are inadequate in that they allow the visibility of heads only and heads do not always carry all relevant information.

¹ The Italian suffix selects also Nouns, but this is not its productive behaviour.

Likewise, theta-grid representations do not work well, and for similar reasons, as Di Sciullo has pointed out, due to the fact that selection based on identity between the theta grid of the verb and the suffix basically reflects a head-to-head relationship. For certain suffixes, however, the thematic information is not enough to allow the formation of all (and only) the possible derivatives. The suffix *-able* which forms adjectives from verbs, for example, selects not only agentive verbs, but also verbs with an experiencer subject and agentives with two internal arguments (theme and goal), requiring therefore a much too complex theta grid, complexity which goes against economy:²

- | | | | |
|-----|-----------|-----------------------------------|-------------------|
| (5) | export | (Agent, Theme) | |
| | detest | (Experiencer, Theme) | |
| | transform | (Agent, Theme, Goal) | |
| | able] A | (Ag, Th ; Exp, Th ; Ag, Th, Goal) | (Di Sciullo 1996) |

The selection made on representations of this sort, moreover, can come up against difficulties across languages: in Dutch, for example, *-baar* ‘-able’ adjectives cannot be followed by the *door* phrase (equivalent to the English *by* phrase) corresponding to the indirect agent, as Booij and van Haften (1988) have shown:

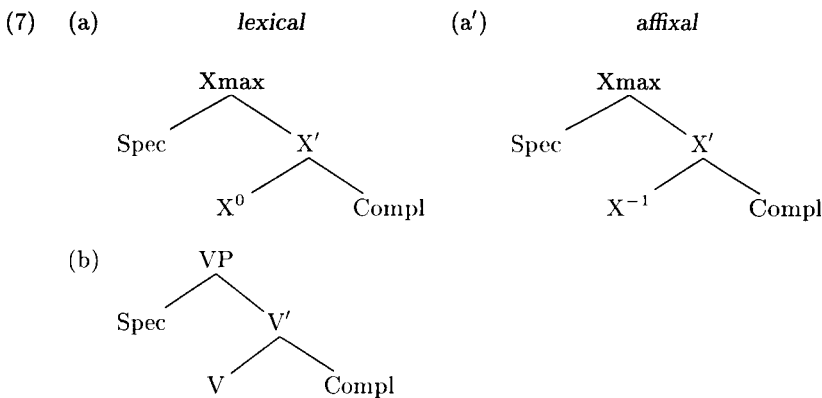
- (6) *dit verschijnsel is verklaarbaar door mij
 ‘this phenomenon is explainable by me’
 *deze soep is eetbaar door mij
 ‘this soup is edible by me’

Consequently, the (Dutch) suffix not having a theta grid matching that of the verb, the selection made by *-baar* ‘-able’ would seem to be unconstrained, or impossible.

² As an anonymous reviewer pointed to us, economy is meant to constrain the theory, not the data; complexity of theta grids, consequently, signals the inadequacy of the theory, but does not render theoretically uneconomical the selection made on theta grids. We think this is correct, but we would like to point out that plurality of theta-grid information in suffix representation misses the point for which selection based on them was advised. Listing of theta grids of all the types of verbs allowing *-able* adjective formation would probably be less “economical” than signalling that the suffix does not select intransitive verbs. Moreover, plurality of theta grids in a suffix representation poses some problems for the U(nitary) O(utput) H(ypothesis) of Scalise (1984). Have we to do with more than one adjectival *-able* suffix? Do *-able* adjectives project different theta grids? Are *-able* adjectives unaccusative contrary to what is suggested by Cinque (1990)?

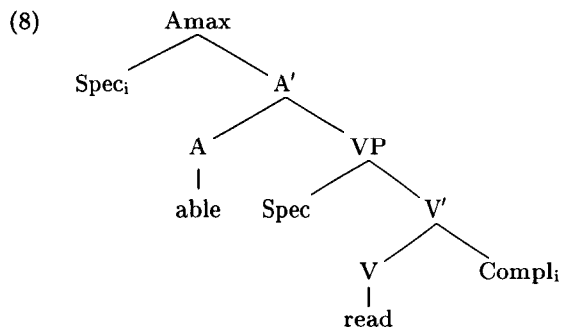
3. X-bar

In recent years, following up on Hale’s (1990) proposals for representing lexical items as an X-bar structure and the “semantic decomposition” of lexical items as an X-bar syntax (cf. Hale–Keyser 1992; 1993), a third proposal was put forth by Di Sciullo, according to which suffixes (i.e., affixal heads), like lexical heads (cf. 7a), have an X-bar representation (cf. 7a’), viz., a representation expressing the projecting capabilities of the suffix. Lexical items, which in derivational word formation constitute the base words, are, on their part, expressed in configurational terms; i.e., the words to which suffixes are attached are represented as an X-bar structure showing the projection of arguments based on their semantic/thematic nature. Such a configuration constitutes what Di Sciullo calls “the complement domain” of the suffix. We illustrate Di Sciullo’s proposal below, where the configuration of an agentive transitive verb is depicted (7b):



The consequence of this is that if affixal heads are considered to project structures in the same way as lexical heads do, affixal and lexical heads become elements having exactly the same status.

The verb in (7b), being a lexical item that can become the base constituent of a derived word, has an X-bar representation that constitutes the complement domain of an affix. As such, it can be put in the Complement position of the suffixal head selecting it. To exemplify the point, let us consider the English adjectival suffix *-able* with its projection structure and its complement domain; according to Di Sciullo’s proposal, what can be obtained is a structure along the following lines.



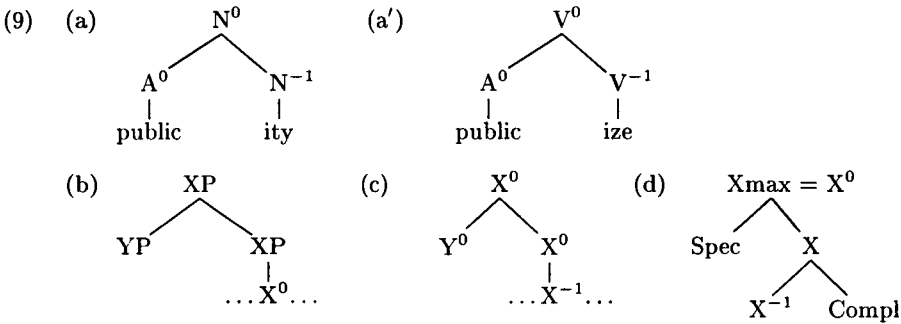
In the upper part of the tree the suffix *-able* with its projecting properties is represented. Its Spec is coindexed with the NP Complement of V. In the lower part, there is the VP which is either the Complement or the complement domain of the suffix at the level of Morphological Form.³ The specific configuration structure of the verbal constituent accounts for the selectional properties of the suffix, indicating that the verbs selected by *-able* are verbs with “thematic” subjects, whatever the theta-role they are assigned.

4. Criticisms

Although we agree in general with Di Sciullo’s criticisms of subcategorization frames and the matching of theta grids as a means of accounting for affixal selection, we have doubts about the functionality of her own proposal. Actually, we do not think that the X-bar structure representation of affixes she proposes, which parallels them with words (i.e., lexical items), is completely convincing. We agree with the claim that affixes (can) carry an argument structure, but we do not see how such an argument structure can be projected in an X-bar configuration. Affixes are bound forms, and as such they will project only when they are part of a word. It is the complex word which projects an X-bar structure, not the affix on its own, even if the arguments are brought about by the affix, as is the case in verb formation from nouns (*canale + izzare* ‘canal + ize’). To simply state that affixes project obscures the possible relation between the argument structure introduced by the affix and that projected by the base word, as is the case with the formation of *-ize* verbs from adjectives (*banale + izzare* ‘trivial + ize’).

³ The level of Morphological Form is, according to Di Sciullo, a level of representation parallel to that of Logical Form.

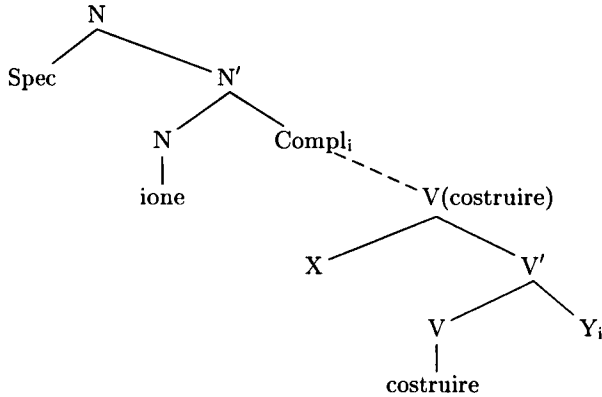
Secondly, one could accept (if with some reservations) the idea put forth by Law (1990), according to whom: (a) the binary branching structure of a derived word (cf. (9a, a')) has the same representation of the elements bearing a head-argument (= complement) relation in an XP adjunction structure (cf. (9b)) and, (b) the representation of the adjunct relation can be extended to words, that is, to X^0 structures, as depicted in (9c) below:



However, we would claim that the correct results do not obtain by maintaining that the base word is the complement domain of the suffix (cf. 9d).

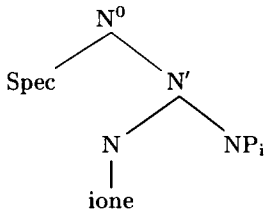
If an affix (a suffix, in the present discussion) is represented as in (9d), the proposal of the “complement domain” of the affix does not work. The Complement projected by the suffix (which would correspond to the Complement projected by the (full) word once the last is formed) does not in fact correspond to the “complement” which constitutes the selectional domain of the affix itself. The Complement position projected by the suffix does not match its own complement domain; i.e., the Complement that the suffix projects (and that will receive saturation in syntax) does not correspond to the configurational representation of the word (the base) selected by the suffix, as can be seen in the following representation where one observes a discrepancy between the Complement of the suffix *-ione* (which is, or, better, will be, a NP) and the “complement domain” of the suffix which is a V(P).

(10)

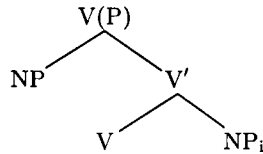


Although it can be said that the NP complement that the suffix projects corresponds to the internal complement of the verbal head (= the base word), there is no matching between the two “complements”. In other words, if the structure is like the following:

(11) (a)



(b)

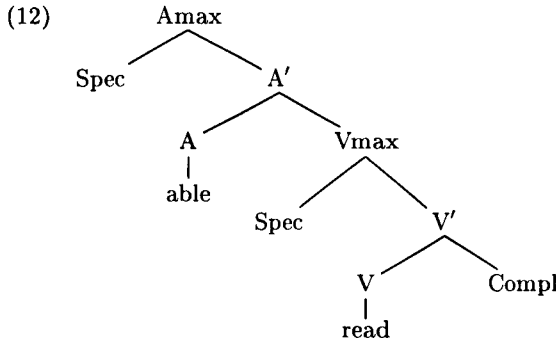


where (11a) is the X-bar projection of the suffix *-ione* and (11b) that of the Verb selected by the suffix *-ione*, what can be said is merely that the complement of V can be co-indexed with the NP complement projected by the suffix.

It seems to us, then, that the two structures cannot be tied together to obtain a “conflation” à la Hale–Keyser (1992; 1993) and a projection structure of the derived word.

In conclusion, it would appear that there is no syntax of suffixed words.⁴

But let us illustrate the point in more detail, taking as an example the X-bar structure proposed for the derivation of *-able* adjectives (already seen in (8) and re-proposed below):



While it is true that: (1) the Spec of the adjective corresponds to the Complement of the verb *read* (*John read the book/the book is readable by John*) and, (2) the verbs having a configuration of the kind illustrated in the above structure are the verbs to which *-able* can be attached (i.e., they are verbs with a “thematic” subject), we cannot say that Vmax is the Complement projected by the suffix. An adjective does not have a Vmax in Complement position. Vmax can be considered as the complement domain of the suffix in a “semantic” sense, but not in structural terms, since an *-able* adjective has no verbal complement.

Thus we can call the (base) words to which suffixes are attached their “complement domain”, but only to the extent to which such an expression is given a particular meaning. In minimalist terms, a complement domain is an internal domain and the “elements of the internal domain are typically internal arguments” of the head (Chomsky 1995, 178). This is not, however, the case in word formation processes.

To sum up, we accept Di Sciullo’s criticism towards selection expressed in terms of subcategorization frameworks and also in part the criticism of theta-grid representations. As for her own proposal, however, though we find the idea that the base word of a derivation process can be represented in configurational terms interesting—not least because configurations are useful means of representing the LCS of a word and affixes often select on an LCS

⁴ For a discussion of the notion of complement in (syntactic) word formation, based on the proposal put forth by Lieber (1992), for example, the reader is referred to Borer (1998).

ground—we have doubts about the proposals that: (1) affixes can have their own X-bar representation and, (2) the domain of selection can be considered to be the Complement of the affixal head.

5. Selection

According to Di Sciullo, one of the points that the parallelism between X^0 and XP structures is based on is the similarity of the mechanisms of selection. Given that mechanisms of selection are crucial both in morphology and syntax, in order to try to maintain that morphology and syntax are separate domains, we would like to illustrate here the different workings of the selection process in morphology and in syntax,⁵ i.e., the different selection lexical heads make with respect to that made by suffixal heads.

It is quite obvious that it is always the head element that operates the selection.⁶ Non-head elements do not select. Therefore: (1) prefixes (when they are not heads) do not select their base and, (2) neither do inflectional affixes select theirs. In other words, only suffixes can operate a selection.

We do not want to discuss either prefixation or inflection here. However, we would point out that the acceptance of a distinction between derivation and inflection has a desired consequence—one which we propose but tentatively, yet in the belief that it fully merits pursuing:

- (13) (a) In a structure [X+Suf] Suf “selects” X
 (b) In a structure [X+Infl] X “selects” Infl

The difference in the selecting constituent is due to the fact that inflection, unlike derivation, is a process of attachment of morphemes bearing “grammatical features”: as such, they are requested by the word X, and this fact renders X the head of the construction. In other words, in an inflectional process it is the nature of X—the word—that imposes its requirements on the affixal part: if X is a noun, the affixal constituent will be a morpheme bearing certain properties characteristic of nouns (gender, number) because nouns (must) carry gender

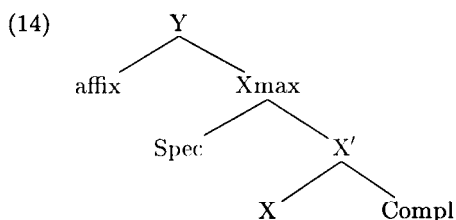
⁵ For other arguments, see Bisetto–Scalise (1997).

⁶ Though what the precise import of such a statement is has not yet been clarified, to say that it is the affixal head that selects its base constituent has important consequences which, however, we will not deal with in the present paper.

and number features.⁷ In derivation processes it is the other way around: in a deverbal nominal, for example, it is not the verb that imposes its requirements on the (derivational) nominal affix, but the other way round.

5.1. The role of the “Complement” in morphology and syntax

At this point, we consider the role of the “Complement” in morphology and syntax. We said above that if the idea that affixal heads have a complement domain of application is accepted, configurational representations of words are the best way of illustrating selection. We saw that the X-bar representation of the structure of a derived word, in which the base word is in the complement position of the suffix, is as follows:



where the base word *X* with its projection structure constitutes the complement domain of the affixal head. However, if the non-head constituent of the word is considered to be the Complement of the affix, a sort of not completely motivated extension of the notion of complement is introduced. Although, as has already been stated, one can accept Law's (1996) proposal on the similarity of *XP* and *X⁰* structures, we do not agree with the statement that in both cases we are faced with complement selection.

In our opinion, affixal selection is actually quite different from lexical selection. First of all, (Law's) complements of suffixes behave in a way that is basically different from that of complements of lexical heads. The latter are, in syntax, the elements which saturate the valency of the head and give rise to an *X* of a higher level. The former are, in morphology, elements that do not

⁷ As pointed out to us by an anonymous reviewer, inflectional affixes are considered to be heads in syntax. This discrepancy between the morphological and the syntactic view of these types of affixes is probably due to the fact that inflectional morphemes are of a syntactic class (the class of functional heads) which plays no role in word formation where a head constituent is intended to be a category changing element. Neither will we discuss this problem in this paper, however.

saturate the valency of the head: the putting together of the complement and the head gives rise to an X of the same (zero) level.

Let us exemplify the case with the structure of a complex word such as the following:

- (15) (a) canalizzare AG, TH 'to canalize'
 / \
 canale izzare AG, TH

- (b) canalizzare un territorio, un campo, etc. 'territory, field'

The (supposed) complement of the suffix *-izzare* (viz. *canale*) does not saturate the argument structure carried by the suffix: the derived word *canalizzare* maintains the argument structure (Agent, Theme) introduced by the suffix. This argument structure will be saturated in syntax only by a true complement, e.g., a NP such as *territorio, campo*, etc. (cf. 15b).

Sometimes, and more importantly, it is the suffix—the head of the construction—that saturates one of the arguments of its complement (i.e., the base word), as is the case with agentive suffixes (e.g., Italian *-tore* or English *-er*), but not *vice versa*. In other words, in word formation the “complement” is the support-constituent of the head, while in syntax it is an element which completes the external valency of the head.

Thus we have to do not only with two different types of “complements”, but also with different mechanisms of selection. Let us illustrate this point by means of two different Italian suffixes, namely *-ato* ‘-ed’ and *-tore* ‘-er’.

5.2. *-ato*

The Italian suffix *-ato* selects verbs and forms adjectives according to a quite elaborate pattern (cf. Bisetto 1994), which is illustrated in (16):

- | | | | | | |
|------------------|-------------|---|--------------|-------------|--------------|
| (16) transitives | illuminare | → | illuminato | 'enlighten' | 'lighted' |
| | trasportare | → | *trasportato | 'carry' | 'carried' |
| intransitives | divorziare | → | divorziato | 'divorce' | 'divorced' |
| | dormire | → | *dormito | 'sleep' | 'slept' |
| psychological | spaventare | → | spaventato | 'frighten' | 'frightened' |
| | contemplare | → | *contemplato | 'admire' | 'admired' |
| ergatives | morire | → | morto | 'die' | 'dead' |
| | venire | → | *venuto | 'come' | 'come' |

As the example illustrates, transitivity, intransitivity, or ergativity, are not decisive notions for the attachment of the suffix: not all transitive verbs, in fact, accept the suffix.

Neither can the selection be expressed on the basis of Aspect, as the following examples illustrate:

| | | | | |
|-------------------------|------|-------------|--------------|------------|
| (17) achievement verbs: | (a) | venire | 'come' | *venuto |
| | | andare | 'go' | *andato |
| | (a') | morire | 'die' | morto |
| | | appassire | 'wither up' | appassito |
| accomplishment verbs: | (b) | celebrare | 'celebrate' | *celebrato |
| | | costruire | 'build' | *costruito |
| | (b') | distruggere | 'destroy' | distrutto |
| | | degenerare | 'degenerate' | degenerato |

It is clear, then, that the suffix *-ato* does not take into account such notions as achievement or accomplishment. The suffix selects the verbs to which it can be attached on the basis of a property that has to do with what happens to the so-called "involved argument". That is to say, the involved argument must acquire a 'change of state' value. The verbs allowing the attachment of *-ato* express a process involving an object characterised by a sort of opposition between an initial state, in which it does not have a particular property, and a final state, in which it does, as the following examples show:

| | | |
|------|------------------------|----------------------------|
| (18) | i fiori sono appassiti | 'the flowers are withered' |
| | Gianni e' divorziato | 'Gianni is divorced' |
| | *Gianni e' corso | lit. 'Gianni is run' |
| | *Gianni e' dormito | lit. 'Gianni is slept' |

As can be observed, the involved argument 'flowers' of the ergative verb *appassire* at the end of the process indicated by the verb is in a status which is different from the initial one.

The same holds for the involved argument of a transitive verb, as (19) illustrates:

| | | |
|------|--------------------------|-------------------------------|
| (19) | l'ombrello e' bagnato | 'the umbrella is wet' |
| | *la favola e' raccontata | lit. 'the fairy tale is told' |

To sum up: the suffix *-ato* does not select its base according to notions such as transitivity or ergativity, nor on the basis of verbal aspect. It selects on the basis of properties (such as 'change of state') that can be considered part of the Lexical Conceptual Structure of the verbs.

5.3. *-tore*

A similar kind of selection is at work with the suffix *-tore* (the equivalent of the English suffix *-er*). This suffix selects the verbs to which it can be attached on the basis of a rather complex mechanism of selection which is illustrated by the following data:

- | | | | |
|--------------------------------------|--------------------------------|--------------------------------|---------------------------------|
| (20) (a') ergative verbs: | *mori-tore, 'die-er' | *arriva-tore, 'arrive-er' | *frana-tore 'crumble-er' |
| (a'') intransitive pronominal verbs: | *ammala-tore, 'get-er sick' | *dispera-tore, 'despair-er' | *assenta-tore 'be-er absent' |
| (a''') psychological verbs: | *preoccupa-tore, 'worry-er' | *disgusta-tore 'disgust-er' | |
| (b) verbs such as <i>temere</i> : | *temi-tore 'fear-er' | | |
| (c) verbs such as <i>sembrare</i> : | *sembra-tore, 'seem-er' | *accadi-tore 'happen-er' | |

In other words, *-tore* does not attach to:

- (21) (i) verbs with non-thematic subject (21a', a'', a'''),
 (ii) verbs with an 'experiencer' subject such as psych. verbs like *temere* (21b)
 (iii) raising verbs (21c)

There are exceptions, obviously, in which we find the suffix attached to psychological verbs of the *temere* class:⁸

- (22) ammira-tore 'admire-er'
 contempla-tore 'contemplate-er'
 ?spaventa-tore 'frighten-er'

⁸ We have accepted here the classification of psychological verbs proposed by Belletti – Rizzi (1988). The theta roles attributed to the subject of the verbs of the *preoccupare* class and the object(s) of the verbs in the *temere* class are differently analysed by Pesetsky (1995). We do not discuss the problem in the present paper and limit ourselves to pointing out that if Pesetsky is right in suggesting that the subject of verbs of the *preoccupare* class is not a Theme (such verbs are then not unaccusative) but a Causer, a different reason from non-thematic subject is needed to explain why such verbs do not form *-tore* agentive nouns, i.e., non-agentivity.

Cases such as these, however, are not counter-examples and can be explained: they are verbs whose semantics includes a notion of 'intentionality', as (23) shows:

- (23) Giovanni ammira deliberatamente i disegni di Miro
 'John deliberately admires Miro's drawings'
 *Giovanni teme deliberatamente l'avanguardia cubista
 'John deliberately fears the cubist avant-garde'

Intentionality is therefore a feature that can render psychological verbs more similar to the verbs to which *-tore* can be attached, i.e., verbs with an agentive or instrumental subject.

Besides agentivity, then, *-tore* requires notions such as intentionality or habitualness, as the following examples, which use both transitive (24a) and intransitive (25b) verbs, show:

- (24) (a) dirottatore 'hijacker'
 costruttore 'builder'
 (b) corridore 'runner'
 bluffatore 'bluffer'

For verbs with a 'non-agentive' subject, another suffix is necessary, for instance, *-ente*:

- (25) dormire 'to sleep' dormiente *dormitore
 soffrire 'to suffer' sofferente *soffritore
 perdere 'to loose' perdente *perditore

Notice, however, that a restriction of this type is not valid across languages; Booij (1986), for example, shows that agentivity in Dutch plays no such role and a word like the following, involving an ergative verb, is well-formed:

- (26) groeier lit. 'grow+er'

Moreover, the Italian suffix *-tore*, like the corresponding English suffix *-er*, cannot be attached to verbs that take a small clause:

- (27) *essere+tore, *consideratore
 *beer, *considerer

An obvious conclusion, then, is that morphological selection cannot be made simply by taking the lexical category into account. Selection appears to be a very refined and delicate mechanism.

As we have seen, a suffix ‘goes into’ the semantics of the word it attaches to, while generally the selection made by an X^0 head, in particular by a verb, can be computed on the basis of its own semantics, and on one’s knowledge of the world. Let us examine, for example, a verb like the following:

(28) trasportare ‘carry’

Such a verb requires that its direct internal complement be characterised by a [–abstract] feature, because only [–abstract] objects can be carried; such a restriction, however, is not always useful, because we know that not every [–abstract] object can be transported. Our knowledge of the world allows us to reject strings such as:

(29) *trasportare un palazzo ‘carry a building’
*trasportare il fiume ‘carry the river’

but to accept strings like those in (30), where the verb is followed by an object carrying the feature [+abstract]:

(30) il vento ha trasportato le sue parole ‘the wind carried away his words’

Such acceptability is due to the metaphorical use of language.

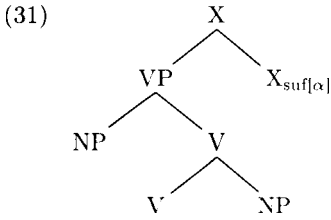
The selection made by X^0 heads, it seems to us, needs a less fine-grained semantics and is less strict than that made by affixal heads.

6. Affix representation

Is there a representation of word formation processes capable of accounting for the restrictions the suffix imposes on the words it selects? Clearly, we think there is, and we would claim that such a representation is in part similar to that proposed by Roeper, and in part to that proposed by Di Sciullo, but with slight modifications. Indeed, we think that there is no need for sets of thematic roles, but only for argument positions and lexical category labels.

It seems to us that affixes cannot be usefully represented with an X-bar structure because they are bound forms and cannot “project” until they are part of an autonomous word, a full word. If an X-bar representation is needed, it can only be used to represent the selected element (when it is a predicate, i.e., a verb or an adjective).

The representation of a suffixal word formation process, capable of pointing to the selecting properties of the suffix, will then be (tentatively) as follows:

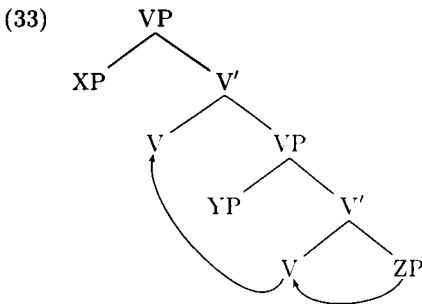


On the left of the suffix X a generic X-bar representation of a predicate is found while on its right (in brackets), there is an α standing for the generic representation of the argument structure carried by the suffix; the arguments of this argument structure are only allowed to be coindexed with the arguments projected by the base word, thus indicating the relationship obtaining between the base word, the suffix and the whole word. To illustrate the point, consider again the case of the two suffixes *-ato* and *-tore*.

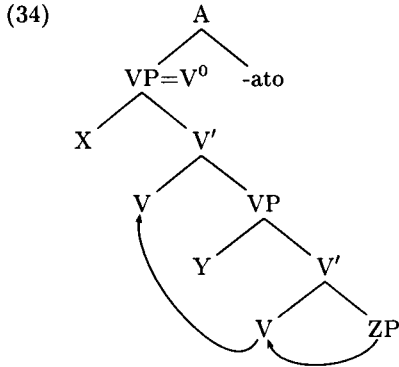
The suffix *-ato*, as we have seen, selects change of state verbs. Limiting ourselves to transitive verbs, we can say that a change of state verb has a ‘causative’ Lexical Conceptual Structure of the following kind:

- (32) X makes Y have the property Z
 e.g., rompere ‘break’ = X makes Y to become broken (Z)

Following the suggestions put forth by Hale–Keyser (1992; 1993), the configurational representation of such an LCS is the one illustrated in (33):



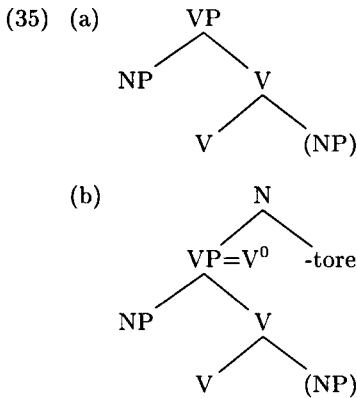
and the representation of the complex adjective will be that illustrated in (34):



The part of the structure that is “visible” on the surface is the binary branching upper part only, while in the “morphological component”, what is relevant is the entire structure.

If a transitive verb has an LCS which does not match the above structure, it cannot be taken as the base word of the derivation process.

Moreover, let us see what happens to our second exemplifying suffix, i.e., *-tore*. This suffix makes its selection depend on the semantic nature of the verbal subject, requiring that it be an agent or an instrument. Assuming that the agent is a thematic external argument, an agentive verb has the representation in (35a), and the word derived with *-tore*, the structure in (35b):



Not having an LCS of this kind, the verbs listed in (20) above (viz., ergatives, psychological verbs, pronominal intransitives and raising verbs) are not selected by the suffix.

7. Conclusions

We will now summarise what we have proposed in this paper:

- (a) an X-bar representation for affixes is problematic because it gives affixes a status of “word” that they do not have;
- (b) the notion of “complement” is necessarily different in morphology and syntax and it is probably not a useful notion in morphology;
- (c) affixal selection is different from lexical selection;
- (d) the representation of affixes is similar to the one proposed by Roeper, except for the fact that we do not make use of theta roles, but only of the argument structure specification introduced by the affix;
- (e) X-bar structures can be adopted to represent the semantics of words in configurational terms;
- (f) the representation we propose is a traditional one, traditional in the sense of “linear”. The X-bar representation we adopted is the representation of the semantics of base words. In our complex word structures, the upper node is then simply an X^0 , rather than a maximal projection in the syntactic sense. As such, it can be put in the traditional position of bases of derived words. Configuration structure becomes simply a way of representing the semantic “decomposition” of a predicate for which an X-bar format is at work. Such a format, however, is relevant for selection only and does not reflect on the derived word. There is no need to consider the configurational domain selected by the affix as a “complement” nor to introduce movement in word formation. Indeed, we think that word formation is a “conflation” *à la* Hale–Keyser only from a semantic point of view. The arguments of the base word can be coindexed with the arguments introduced by the affix, but they do not “conflate” in the meaning of the final word.
- (g) From the above considerations it follows that syntax cannot absorb morphology, since morphology cannot adopt the syntactic machinery without relevant adjustments.

References

- Aronoff, M. 1976. *Word formation in generative grammar*. The MIT Press, Cambridge MA.
- Belletti, A. – Rizzi, L. 1988. Psych-verbs and Theta-Theory. In: *Natural Language and Linguistic Theory* 6: 291–325.
- Bisetto, A. 1994. Gli aggettivi in *-(x)to*. In: Borgato, G. (ed.): *Teoria del Linguaggio e Analisi Linguistica, XX Incontro di Grammatica Generativa*, 63–81. Unipress, Padova.

- Bisetto, A. – Scalise, S. 1997. L'autonomie de la morphologie. XVI Congrès International des Linguistes, Paris, Elsevier, CD ROM.
- Booij, G. 1986. Form and meaning in morphology: the case of Dutch 'Agent Nouns'. In: *Linguistics* 24: 503–18.
- Booij, G. – van Haften, T. 1988. The external syntax of derived words: evidence from Dutch. In: *Yearbook of Morphology* 1: 29–44.
- Borer, H. 1998. Morphology and syntax. In: Spencer, A. – Zwicky, A. (eds): *Handbook of morphology*, 151–90. Blackwell, Cambridge MA & Oxford.
- Chomsky, N. 1965. *Aspects of the theory of syntax*. The MIT Press, Cambridge MA.
- Chomsky, N. 1995. *The Minimalist Program*. The MIT Press, Cambridge MA.
- Cinque, G. 1990. *Types of A-bar dependencies*. The MIT Press, Cambridge MA.
- Di Sciullo, A.M. 1992. Selection and derivational affixes. Paper presented at the Fifth International Morphology Meeting in Krems.
- Di Sciullo, A.M. 1993. The complement domain of a head al morphological form. In: *Probus* 5: 95–126.
- Di Sciullo, A.M. 1996. Atomicity and relatedness in configurational morphology. In: Di Sciullo, A.M. (ed.): *Configurations essays on structure and interpretation*, 17–41. Cascadilla Press, Somerville.
- Gruber, J. 1967. *Studies in lexical relations*. Ph.D. dissertation, MIT, Cambridge MA.
- Hale, K. 1990. *The syntax of lexical word formation*. Ms., MIT, Cambridge MA.
- Hale, K. – Keyser, J. 1992. The syntactic character of thematic structure. In: Roca, I.M. (ed.): *Thematic structure. Its role in grammar*, 107–41. Foris, Berlin–New York.
- Hale, K. – Keyser, J. 1993. *The view from Building 20. Essays in linguistics in honor of Sylvain Bromberger*. The MIT Press, Cambridge MA.
- Jackendoff, R. 1972. *Semantic interpretation in generative grammar*. The MIT Press, Cambridge MA.
- Law, P. 1990. Heads, arguments and adjuncts in Derivational Morphology. In: *MIT Working Papers in Linguistics* 12. The MIT Press, Cambridge MA.
- Law, P. 1996. On some syntactic properties of word-structure and modular grammars. In: Di Sciullo, A.M. (ed.): *Projections and interface conditions*, 28–51. Oxford University Press, Oxford.
- Lieber, R. 1980. *On the organization of the lexicon*. Ph.D. dissertation, IULC, Bloomington, Indiana.
- Lieber, R. 1983. Argument linking and compounds in English. In: *Linguistic Inquiry* 14: 251–85.
- Lieber, R. 1992. *Deconstructing morphology*. The University of Chicago Press, Chicago.
- Pesetsky, D. 1995. *Zero syntax: experiencers and cascades*. The MIT Press, Cambridge MA.

Roeper, T. 1987. Implicit argument and the head-complement relation. In: *Linguistic Inquiry* 18: 267-310.

Scalise, S. 1984. *Generative morphology*. Foris, Dordrecht.

Address of the authors: Antonietta Bisetto and Sergio Scalise
Department of Foreign Languages
University of Bologna
Via Cartoleria 5
40124 Bologna
Italy
bisetto@lingue.unibo.it
scalise@lingue.unibo.it

RESULTATIVE ADJECTIVES IN POLISH*

BOŻENA CETNAROWSKA

Abstract

This paper argues that resultative *-ł*-adjectives in Polish, such as *zbiegły* ‘escaped’ and *zwiędły* ‘faded’, are derived from past (*-ł*-stem) participles by means of conversion (or zero-derivation). Evidence is presented in favour of the lexical (rather than the syntactic) character of the derivation, in the course of which resultative adjectives inherit the event structure of verbal bases. It is suggested that the availability of a resultative *-ł*-adjective is indicative of the unaccusative status of the related verb in Polish.

1. Introduction

The main aim of this paper is to argue that Polish resultative adjectives terminating in the sequence *-ły* (such as *pożółkły* ‘that has become yellow’, *zbiegły* ‘escaped’ or *zwiędły* ‘withered’) are derived by means of conversion (interpreted as an affixless lexical operation) from inflectional past participles. In the course of the discussion I will also present evidence for the validity of the Unaccusativity Hypothesis in Polish.

The term “resultative” is defined in Nedjalkov–Jaxontov (1988, 6) as “applied to those verb forms that express a state implying a previous event”. Traditionally these forms are referred to as “participles” when they belong to

* This research was supported in its initial stage by a grant from the State Committee for Scientific Research (KBN), nr 1 H01D 00511. I was able to continue the research during my stay as a Fulbright scholar at the University of Massachusetts in Amherst (USA). A preliminary version of this paper was presented at the 8th International Morphology Meeting in Budapest in June 1998. I would like to thank the participants at the conference as well as Martin Haspelmath, Angeliek van Hout, Ferenc Kiefer, Tom Roeper and Peggy Speas for discussing with me parts of the material included here. Christina Bethin, Kyle Johnson, Bożena Rozwadowska, and an anonymous reviewer provided detailed comments on an earlier version of the paper. I am particularly grateful to Robert Rothstein for reading very carefully through the manuscript and making valuable suggestions aimed at improving the final shape of this work. All errors are my own.

the inflectional paradigm of the verb (i.e., when their occurrence for each verb is predictable) and as “deverbal adjectives” if they are outside the inflectional paradigm of related verbs. This terminological distinction is not, however, adhered to strictly in the linguistic literature (see, for instance, Haspelmath 1994). As pointed out in Haspelmath (1994, 159), resultative deverbal adjectives and participles can be oriented either towards the object of transitive verbs or towards the subject of intransitive verbs. If they are object-oriented they are conventionally termed “passive participles/adjectives”, as in the case of *broken heart* and *stolen money* in English. Resultative forms in English which are subject-oriented, such as *fallen leaf* and *escaped convict*, are referred to as “adjectival past participles” in Ackerman–Goldberg (1996) (while in Levin–Rappaport (1986) they are termed “adjectival passives”). When talking about subject-oriented forms in Polish I will use the terms “non-passive resultative adjectives”, “subject-oriented resultative adjectives” or “active resultative adjectives” interchangeably.

I will focus here on non-passive resultative adjectives in Polish which terminate in the sequence *-ły* (where *-t-* can be regarded either as a stem-forming or a derivational suffix and *-y* is an inflectional ending, marking ‘nominative-sg-masc’), as exemplified in (1):

- | | | | | |
|-----|-----------|-----------|-----------|-----------------------|
| (1) | dojrzały | ‘ripe’ | wychudły | ‘thin, emaciated’ |
| | przybyły | ‘arrived’ | zgniły | ‘putrid, rotten’ |
| | upadły | ‘fallen’ | zbiegły | ‘escaped’ |
| | posiwiały | ‘grey’ | zdziczały | ‘that has grown wild’ |
| | umarły | ‘dead’ | zwiędły | ‘faded, withered’ |

Less attention will be given to subject-oriented resultative adjectives which contain the suffix *-n-/-t-*,¹ as in (2).

- | | | | | |
|-----|--------------|-------------|------------|------------------|
| (2) | oburzony | ‘indignant’ | wypoczęty | ‘rested’ |
| | popękany | ‘cracked’ | schylony | ‘stooping’ |
| | uśmiechnięty | ‘smiling’ | zmarznięty | ‘frozen, chilly’ |

As shown in Section 2, many deverbal adjectives/participles with the suffix *-n-/-t-* can be oriented either towards the subject or object of a clause. They are thus ambiguous between passive and non-passive interpretation.

¹ The occurrence of the *-n-* or *-t-* variant of the morpheme in question is predictable on the basis of the phonological make-up of the verb stem. The *-t-* allomorph is selected when the verb stem ends in a consonant in its underlying phonological representation, as in *myty* ‘washed, clean’ (from *myć* ‘to take’), *spuchnięty* ‘swollen’ (from *spuchnąć* ‘to swell’) or *tarty* ‘ground’ (from *trzeć* ‘to grind’). The *-n-* allomorph occurs elsewhere. See, for instance, Rothstein (1970) for more discussion.

Section 2 will consider the issue of whether the derivation of non-passive resultative adjectives can take place in the syntax (as has been proposed for Dutch deverbal adjectives and participles in Drijkoningen 1992). I will highlight the idiosyncrasy involved in the derivation of Polish *-t-* adjectives, both in their form and in their semantic interpretation.

In sections 3 and 4 I will discuss ways of deriving *-t-* adjectives in the lexicon. The assumption prevalent in current Polish word-formation studies and in academic coursebooks on Polish morphology is that such adjectives are formed from verb stems by means of the derivational suffix *-t-* (see Grzegorzczkowska 1979; Kallas 1984). In section 3 I will demonstrate that resultative *-t-* adjectives exhibit stem allomorphy of the same type as finite past tense forms, which is not predicted by the hypothesis that *-t-* attaches to verb stems. A proposal I will consider briefly in section 3 is the derivation of *-t-* adjectives through conversion from the so-called “third-degree stems” of related verbs (where the term “third-degree stem” denotes one of several canonical shapes of each verb, as in Laskowski 1975 and Szymanek 1985). I will indicate some disadvantages of this proposal and endorse the next hypothesis, which states that *-t-* adjectives are derived through conversion (adjectivization) from past participles of related verbs. In section 4 I will comment on the diachronic and synchronic status of past participles in Polish. Then I will offer an account for why non-passive resultative adjectives can occur with the kind of modifiers and adjuncts that are compatible with finite verb forms. The account will employ the notion of the event structure (as postulated in, among others, Pustejovsky 1991 or Grimshaw 1990).

In order to provide additional support for the hypothesis that *-t-* adjectives are derived from inflectional past participles through adjectivization, I will present in section 6 cross-linguistic evidence and argue for the existence of the unaccusative/unergative distinction in Polish verbs.

2. The hypothesis of syntactic derivation of non-passive resultative adjectives

2.1. Internal syntax of phrases headed by resultative adjectives

Babby (1973) proposed syntactic derivation of active participles in Russian within the framework of transformational grammar to account for the occurrence of tense, aspect and verb complementation in nonfinite verb forms.²

Adopting more recent versions of the generative grammar framework, Schoorlemmer (1995) argues for Russian and Drijkoningen (1992) and van der Putten (1997) propose for Dutch that the derivation of deverbal adjectives/participles can take place in syntax.

Drijkoningen postulates for Dutch that the derivation of present, perfective and passive participles involves adjunction of a category-changing (adjectivizing) affix to an appropriate verbal projection (V' or V^{\max}). As a consequence of their derivation in the syntax, participles/deverbal adjectives are expected to show the internal syntax of finite verb forms, for instance allowing the same set of arguments and modifiers, in spite of having the external distribution and inflectional properties of adjectives.

The internal syntax of many phrases headed by non-passive resultative adjectives in Polish—either containing the suffix *-t-* or *-n-/-t-*—resembles the internal syntax of phrases containing finite verb forms and passive/present participles in many respects. Non-passive resultative adjectives in (3a–c) can take past tense temporal adverbials, similarly to the passive participle in (3d) and the present participle in (3e).

- (3) (a) przybyły przed chwilą posłaniec ‘a messenger that arrived a moment ago’
 (b) zgnile zimą ziemniaki ‘potatoes that grew rotten in winter’
 (c) rozkwitłe dziś rano tulipany ‘tulips that have opened this morning’
 (d) zbita przed chwilą przez ciebie szyba ‘the window pane that was broken by you a moment ago’
 (e) płaczący wczoraj chłopiec ‘the boy who was crying yesterday’

Premodifying non-passive resultative adjectives can also occur with resultative phrases, locative (goal or source) expressions, instrumental adjuncts and various prepositional complements, as illustrated in (4).

² Babby (1998) adopts the position that the formation of Russian deverbal adjectives is performed by a lexical rule.

- (4) (a) przemoknięty/przemokły do suchej nitki chłopiec 'a boy who is drenched to the skin'
(b) przerdzewiały do fundamentów most 'a bridge that is rust-eaten, down to the foundations'
(c) zeschnięte/zeschnię na twardą skorupę błoto 'mud that has dried until it formed a hard crust'
(d) zamarzył na śmierć żołnierz 'a soldier who froze to death'
(e) zmarznięte na kość mięso 'meat that is frozen to the marrow'
(f) spróchniałe do korzeni drzewo 'a tree that is rotten to the roots'
(g) przybyły na miejsce tragedii prezydent 'the president who has arrived at the scene of the tragedy'
(h) opadły na ziemię liść 'a leaf that has fallen to the ground'
(i) przybyły z Gdańska mężczyzna 'the man who has come from Gdańsk'
(j) porosły mchem kamień 'a moss-covered stone'
(k) pozostałe po obiedzie resztki 'left-overs remaining after dinner' (lit. remained after dinner left-overs)
(l) przepadłe bez wieści dzieci 'children who have disappeared without a trace'
(m) zetlałe na węgiel belki 'beams that have smoldered away, turning into charcoal'
(n) zmarły na zapalenie płuc kuzyn 'a cousin who died because of pneumonia'
(o) rozgorzały na nowo spór 'the quarrel that has broken out again'

An explanation for the data in (3)–(4) will be offered in section 5.

The issue of the division of labour between syntax and the lexical component is highly controversial. However, morphological processes operating in the syntax are commonly expected to have properties of canonical inflection,

such as predictability of occurrence, semantic compositionality and formal regularity of the output (see e.g., Anderson 1982 and Scalise 1988).³

In section 2.2 below I will show that the derivation of non-passive resultative adjectives is not fully productive and involves a high degree of idiosyncrasy, in both the form of the suffix and the semantic interpretation of the derived formation.

2.2. The distribution of the *-t-* and *-n-/-t-* affix

As was illustrated in (1)–(2) in section 1, non-passive resultative adjectives in Polish fall into two large classes: those containing the morpheme *-t-* and those exhibiting the morpheme *-n-/-t-*.

It is important to note that the morpheme *-n-/-t-* also appears in passive participles, e.g., *kochany* ‘loved’, *poszukiwany* ‘looked for, wanted’ or *zamordowany* ‘murdered’. Thus in some cases a single morphological shape is associated both with the passive participle of a transitive verb and the non-passive resultative adjective from the corresponding intransitive (“reflexive”) verb, as in (5).⁴

³ A different view of the distinction between syntactic and lexical morphology is presented in van der Putten (1997). Van der Putten (1997, 34ff) argues that syntactic morphology makes use of syntactic vocabulary, referring to notions such as ‘subject’ and ‘object’. Restrictions on processes of syntactic derivation should be stated in syntactic terms. Lexical morphology uses the vocabulary of the lexicon, making reference to thematic notions and semantic features, such as ‘animate’, ‘human’ etc. However, he adopts the assumption that syntactic morphological processes are fully productive while lexical processes are semi-productive or unproductive. Yet another approach towards the relation between syntax and morphology is postulated within the model of Parallel Morphology, as illustrated in Schoorlemmer (1995). Within this model certain morphological processes can occur both presyntactically (lexically) and in the syntactic component. Within the model of Distributed Morphology, as presented in Marantz (1999), derivation of words is assumed to take place in the syntactic component. The distinction between lexical and syntactic derivation, adopted for instance in Anderson (1982), is restated in Marantz (1999) as a difference between derivation from roots and derivation from ‘little x’, respectively.

⁴ Assuming the separation of rules into rules of semantic derivation (L-derivation) and morphological affixation (M-derivation), as in Beard (1995) or Szymanek (1985), we can say that the multifunctional morpheme *-n-/-t-* is associated both with the rule forming non-passive resultative adjectives and with the rule deriving passive participles (see Spencer 1991 for more discussion of the Separation Hypothesis). Both rules give rise to adjectives which are able to modify the internal argument of the verb.

- (5) (a) przestraszony 'frightened' (cf. przestraszyć się 'to get frightened', przestraszyć 'to frighten (sb)')
- (b) zбитy 'broken' (cf. zbić się 'to get broken, to break by itself', zbić 'to break sth')
- (c) otwarty 'open; opened' (cf. otworzyć się 'to open by itself', otworzyć 'to open sth')
- (d) zamknięty 'close, closed' (cf. zamknąć się 'to close by itself', zamknąć 'to close sth')
- (e) ogolony 'shaved, shaven' (cf. ogolić się 'to shave oneself', ogolić 'to shave sb')
- (f) umyty 'washed' (cf. umyć się 'to get washed', umyć 'to wash sth/sb')
- (g) zgromadzony 'gathered' (cf. zgromadzić się 'to assemble, meet', zgromadzić 'to gather sb/sth')

The form *zбитy* 'broken' in (5b) requires a passive interpretation when it implies an event involving two participants (the Agent and the Theme/Patient), as denoted by the sentence *Piotr zbił dziś dwie szklanki* 'Peter broke two glasses today'. If interpreted as referring to a one-participant event (involving the Theme but no external Agent/Causer, as in *Szklanka się zbita* 'The glass broke'), the participle *zбитy* 'broken' receives a non-passive resultative interpretation.

The *-n-/-t-* morpheme is the normal marker of resultative adjectives from all types of reflexive verbs, including the reflexiva tantum in (6), i.e., verbs which are obligatorily reflexive.

- (6) (a) spóźniony 'late' (from spóźnić się 'to come late')
- (b) uśmiechnięty 'smiling' (from uśmiechnąć się 'to smile')
- (c) zamyślony 'lost in thought' (from zamyślić się 'to fall into thought')
- (d) spocony 'sweaty' (from spocić się 'to sweat')
- (e) zaziębiony 'suffering from a cold' (from zaziębić się 'to catch cold')

The pattern of distribution of the suffixes *-n-/-t-* and *-t-* in non-passive resultative adjectives is, however, far from being neat.

There are a number of instances where parallel *-t-* and *-n-/-t-* forms exist. Although intransitive verbs are expected to derive resultative *-t-* adjectives, the forms with *-n-/-t-* are more frequently used than the *-t-* adjectives in the majority of the examples in (7).

- (7) (a) zamaznięty or zamazły 'frozen' (from zamaznąć 'to freeze')
- (b) przesiąknięty or przesiąkły 'permeated (e.g., with sweat)' (from przesiąknąć 'to become permeated')
- (c) nasiąknięty or nasiąkły 'that has absorbed (water)' (from nasiąknąć 'to absorb')
- (d) namoknięty or namokły 'saturated (with water)' (from namoknąć 'to become wet')
- (e) spierzchnięty or spierzchły 'chapped (of skin)' (from spierzchnąć 'to become chapped')
- (f) wyschnięty or wyschły 'dried' (from wyschnąć 'to dry')
- (g) spuchnięty or spuchły (rare) 'swollen' (from spuchnąć 'to swell')
- (h) zziębnięty or zziębły (rare) 'freezing, chilled, that feels cold' (from zziębnąć 'to start feeling chilly')
- (i) zachrypnięty or zachryply 'that has become hoarse' (from zachrypnąć 'to become hoarse')

The infinitives of the verbal bases for parallel *-n-/-t-* and *-t-* adjectives in (7) contain the suffix *-nq-*. This suffix is lost (truncated) in front of the morpheme *-t-*. The resultative adjectives derived with the *-n-/-t-* suffix preserve the stem-final *-nq-* morpheme in (7) and bear greater resemblance to the infinitive and to the non-past finite verb forms than *-t-* adjectives. Consequently, they are preferred over *-t-* adjectives, even though their derivation violates the principle of reflexive bases⁵ for *-n-/-t-* non-passive resultative adjectives. A similar explanation is not available, however, for the occurrence of *-n-/-t-* resultative adjectives derived from the intransitive verbs in (8).

- (8) (a) umarty (dial.) or umarły 'dead' (from umrzeć 'to die')
- (b) zemdlony or zemdlący (rare) 'in a faint' (from zemdleć 'to faint')
- (c) znieruchomiony (rare) or znieruchomiały 'motionless, that has ceased moving' (from znieruchomieć 'to become motionless')
- (d) skwaśnięty (dial.) or skwaśniały 'sour' (from skwaśnieć/skwasić się 'to turn sour')
- (e) zgnity (dial.) or zgnily 'rotten, putrid' (from zgnić 'to become rotten')
- (f) popękany or popękaly (rare) 'cracked' (from popękać 'to crack')

⁵ This principle was proposed in section 2.2 on the basis of the data in (5)–(6).

Some intransitive verbs (such as those in (9)) do not derive *-t-* adjectives. They give rise to *-n-/-t-* adjectives.

- (9) (a) *zwariowany* 'that has gone mad' (from *zwariować* 'to go mad')
 (b) *pęknięty* 'cracked' (from *pęknąć* 'to burst, to crack')
 (c) *padnięty* (coll.) 'dead tired' (from *paść* 'to fall')
 (d) *nawalony* (coll.) 'broken; drunk' (from *nawalić* 'to go wrong, to break down' or *nawalić się* 'to get drunk')
 (e) *wypoczęty* 'rested' (from *wypocząć* 'to rest')

The facts presented in (7)–(9) may be construed as testifying to the diminishing productivity of *-t-*, which is giving ground to the suffix *-n-/-t-*.

The suffix *-t-* is, however, not yet completely moribund. It appears in resultative adjectives formed fairly recently from highly colloquial (or slang) terms in (10a–d).

- (10) (a) *zgreździały* (coll.) 'that has become old' (cf. *zgreździeć* 'to become old and gaga')
 (b) *zdziaździały* (coll.) 'that has become old' (cf. *zdziaździeć* 'to become old and gaga')
 (c) *ocipiały* (vulg.) 'that has gone mad' (cf. *ocipieć* (vulg.) 'to go mad')
 (d) *zapyziały* (coll.) 'sloppy, untidy, uncared-for' (cf. *zapyzieć* (coll.) 'to become untidy and sloppy')

To complete the picture of the distribution of *-t-* and *-n-/-t-* suffixes in non-passive resultative adjectives, let us consider the forms in (11). The reflexive verbs in (11) derive *-t-* adjectives, instead of the expected formations with the *-n-/-t-* morpheme.

- (11) (a) *zsiadłe mleko* 'sour milk' (from *zsiąść się* 'to become sour and curdled')
 (b) *przełękłe dziecko* 'frightened child' (from *przeleknąć się* 'to become frightened')
 (c) *postarzały mężczyzna* 'man that has grown old' (from *postarzeć się* 'to grow older')
 (d) *rozeschłe drzewo* 'wood that has dried up' (from *rozeschnąć się* 'to dry and crack')
 (e) *zleżały towar* 'goods that have become spoiled by lying in the shop for too long' (from *zleżeć się* 'to become shopworn, to lie for too long in the shop')

One can presumably attribute the occurrence of the unexpected *-t-* or *-n-/-t-* adjectives to the influence of analogy. The intransitive verb *uschnąć* 'to wither'

forms the resultative adjective *uschły* 'withered'. Consequently, the related reflexive verb *zeschnąć się* 'to dry apart' derives the adjective *rozeschły* 'that has dried apart' in (11d) instead of the adjective *rozeschnięty*. The influence of analogy is observable in lexical—and not syntactic—derivation (as noted in van der Putten 1997). Since the choice of the *-n-*/*-t-* or *-t-* suffix is not fully predictable in (11) and in (7)–(9), such resultative adjectives should be regarded as listed (by virtue of showing formal idiosyncrasies).

2.3. Semantic idiosyncrasies

Resultative *-t-* adjectives occasionally develop idiosyncratic readings. For instance, they may require the interpretation characteristic of active present *-qc-* participles, as in the case of the adjectives mentioned in (12).

- (12) (a) *czuły* 'sensitive, caring' (from *czuć* 'to feel')
- (b) *dbały* 'careful' (from *dbać* 'to take care')
- (c) *nawisły* 'overhanging' (from *nawisnąć* 'to overhang')
- (d) *przyległy* 'adjoining' (from *przylec* 'to adjoin')
- (e) *trwały* 'constant, durable' (from *trwać* 'to last, to persist')
- (f) *wytrwały* 'persistent, persisting' (from *wytrwać* 'to persist')
- (g) *zamieszkały* 'residing, resident' (from *zamieszkać* 'to reside, to live')

Secondly, *-t(y)* adjectives may show passive and/or potential interpretation, as in (13).

- (13) (a) *zamieszkały dom* 'a house that is lived in'
- (b) *niedościgły* (rare) 'that was not reached'
- (c) *niedosięgly* 'unattainable, out of reach'
- (d) *obsiadły* 'occupied, filled, taken up'

The meaning of *-t(y)* adjectives may depart considerably from the meaning of related verbs, as illustrated in (14).⁶

⁶ It is worth noting that the majority of *-t-* adjectives which show semantic idiosyncrasies are related to non-prefixed verbs and/or non-durative (atelic) verbs. Resultative adjectives derived from telic change-of-state verbs and from "degree achievement" verbs are fairly regular in their semantics.

- (14) (a) *bywały* (w świecie) 'experienced, knowledgeable' (cf. *bywać* 'to frequent')
- (b) *rosły* 'tall' (cf. *rosnąć* 'to grow')
- (c) *przyszły* 'future' (cf. *przyjść* 'to come')
- (d) *wzniosły* 'lofty; noble' (cf. *wznieść się* 'to rise, to soar')
- (e) *przeciągły* 'protracted, lengthy' (cf. *przeciągnąć (się)* 'to lengthen, to last')

The internal structure of *-t-* adjectives may become totally opaque, as in the case of *smukły* 'slender' (which has no related verb **smuknąć* 'to become slender') and other forms in (15).

- (15) (a) *pociągły* 'oblong, slender' (cf. *pociągnąć* 'to pull')
- (b) *przebiegły* 'shrewd, cunning' (cf. *przebiec* 'to run (a certain distance)')
- (c) *smagły* 'having a dark complexion' (cf. *smażyć* 'to fry')
- (d) *okazały* 'magnificent' (cf. *okazać się* 'to appear, to turn up')
- (e) *wypukły* 'convex' (cf. *wypuklić* (rare) 'to belly sth out')

Furthermore, there are a number of potential well-formed *-t-* adjectives which are not institutionalized, are marked as dialectal or have fallen out of use, e.g., the non-occurring forms *?*przytyły*, *?*utyły* 'that has grown fatter' (potentially formed from *przytyć* 'to become slightly fat' and *utyć* 'to become fat') or *?*znikły* 'that has disappeared' (from *zniknąć* 'to disappear') and *?*uciekły* 'escaped' (from *uciec* 'to escape'). Bartnicka (1970) and *SJP* mention a number of *-t-* (and *-n-/-t-*) formations that are archaic, dialectal or function as occasional forms, e.g., *?popękały* 'cracked' (cf. established *popękany* 'cracked'), *rozpękły* (arch.) 'burst' (cf. *pęknięty* 'burst'), *najadły* (dial.) 'full' (cf. established *najedzony* 'full, that has satisfied one's appetite'), *nieudały* (arch.) 'failed' (cf. established *nieudany* 'failed') and *?przeminięły* 'elapsed' (cf. *miniony* 'past'). The occurrence of such 'accidental gaps' is a feature of a lexical (and not a syntactic) process.

Lexical processes are not fully productive but they can show a fair degree of productivity. Speakers of Polish are able to contrast resultative adjectives in *-t-(y)* and in *-n-/-t-(y)*, (derived from cognate intransitive and transitive verbs) as active and passive, respectively.

- (16) (a) *poczerniały* 'that has grown slightly black' vs. *poczerniony* 'that has been made slightly black' (cf. *poczernieć* 'to become black' and *poczernić* 'to blacken sb/sth')
- (b) *ocalały z pożaru* 'that has survived a fire' vs. *ocalony z pożaru* 'that has been saved from a fire' (cf. *ocaleć* 'to survive, to escape' and *ocalić* 'to save sb/sth')
- (c) *wychudły* 'that has grown thin' vs. *wychudzony* 'that has been starved' (cf. *wychudnąć* 'to become thin' and *wychudzić* 'to make lean')
- (d) *rozmiękły* 'that has gone soaked and soft' vs. *rozmiękczony* 'that has been soaked and softened' (cf. *rozmięknąć* 'to become soaked and soft' and *rozmiękczyć* 'to soak and make soft')
- (e) *zamieszkały* 'resident' vs. *zamieszkaný* 'that has been lived in' (cf. *zamieszkać* 'to come to live, to occupy (as a residence)')
- (f) *postarzały* 'that has grown older' vs. *postarzony* 'that has been made to look older' (cf. *postarzeć się* 'to grow old' and *postarzyć* 'to make look older')

Such a contrast can be made even in the cases when the *-n-/-t-* or *-t-* form is ambiguous in non-contrastive contexts. In (16c), for instance, the form *wychudzony* 'that has been starved' optionally allows the non-passive reading 'that has starved oneself'. Nevertheless, the passive interpretation of *wychudzony* is obligatory in the phrase *nie wychudły lecz wychudzony* 'not one that has grown thin by his own will but one made to grow thin'. The adjective *zamieszkały* 'resident' in (16e) occasionally occurs in the passive sense, as in the phrase *zamieszkały dom* 'a house that has been lived in'. However, this adjective is recognizable as the non-passive (subject-oriented) form when contrasted with the passive resultative form *zamieszkaný* 'that has been lived in'. The data in (16) suggest that the derivational patterns for the formation of passive and non-passive resultative adjectives are fairly robust in Polish.

The facts presented in this section support, moreover, the hypothesis that the derivation of non-passive resultative adjectives should take place in the lexicon since the morphological shape of derived adjectives (i.e., the choice between *-n-/-t-* and *-t-* forms) and their acceptability cannot be fully predicted.

In sections 3 and 4 I will investigate what type of lexical process gives rise to the occurrence of deverbal *-t-* adjectives and which verbal forms function as their derivational bases.

3. Derivatives from third-degree verb stems

The discussion of formation of non-passive resultative *-t-* adjectives in the immediately preceding section was based on the tacit assumption (adopted

in the majority of current Polish morphology textbooks) that the *-t-* suffix attaches—either in the course of lexical or syntactic derivation—to basic verb stems. Now I will consider the advantages and disadvantages of an alternative analysis, according to which the morpheme *-t-* in adjectives such as *zwiędły* ‘withered’ is not an adjective-forming derivational suffix but part of the third-degree verb stem.

Such an analysis employs the distinction between various stems of Polish verbs postulated, for instance, in Laskowski (1984). Each verb has its basic stem and secondary stems which are derived from the basic stem by means of stem-forming morphemes, such as *-t-*, *-n-/-t-*, *-qc-* and *-∅-/-ij-*:

- (17) (a) *biegać* ‘to run’
 bieg- -a- -ć
 basic stem + thematic vowel + infinitival infl. ending
- (b) the secondary past tense stem:
 bieg- -a- -ł-
 basic stem + thematic vowel + stem-forming suffix
- (c) *biegałam* ‘ran-1sg.fem’
 biegal- -am
 secondary stem + number/gender/person infl. ending

Laskowski expresses the conviction that stem-forming morphemes do not have any clear semantic or grammatical meaning. The morpheme *-t-* is not, according to him, an exponent of the past, nor is the morpheme *-n-/-t-* regarded as the marker of passivity. As shown in (18), the secondary stem in *-t-* (which is called the past tense stem or the third-degree stem) occurs not only in past-tense forms, but also in subjunctive forms and analytical future-tense forms. The past-tense stem is also the basis for the formation of the anticipatory *-łszy/-wszy* participle, or verbal adverb, which occurs in the inflectional paradigm of perfective verbs, e.g., *zobaczywszy* ‘having seen’, *przebiegłszy* ‘having run’:

- (18) (a) *widziałam* ‘I-fem.-saw’
widziałabym ‘I-fem. would see’
będę widziała ‘I-fem. will see’
- (b) *słyszał* ‘he-heard’
słyszałby ‘he would hear’
będzie słyszał ‘he will hear’
- (c) *biegły* ‘they-fem. ran’
biegłyby ‘they-fem. would run’
będą biegły ‘they-fem. will run’

The secondary stem containing the stem-forming affix *-n-/-t-* similarly shows a wide range of uses. It occurs in passive participles, in verbal nouns and in the impersonal past tense forms, as illustrated in (19) below:

- | | | | | | |
|----------|----------|----------------|-----|-------|---------------|
| (19) (a) | kochany | 'loved' | (b) | bity | 'beaten' |
| | kochanie | 'loving' | | bicie | 'beating' |
| | kochano | 'one loved sb' | | bito | 'one beat sb' |

The analyses carried out in Laskowski (1975; 1984) are compatible with the approach advocated in Aronoff (1994) and Sadler *et al.* (1997), where it is argued that verb stems are not uniquely characterized in terms of morphosyntactic features (such as present/past or active/passive). Consequently, a single verb stem can give rise to forms which bear opposite morphosyntactic features, such as the future active participle and the past passive participle in Latin.

The assumption that resultative *-t-* adjectives are regularly formed from *-t-* secondary stems accounts for the formal identity of those adjectives and stems of the past tense forms.⁷ For the verbs given in (20a–c), the resultative *-t-* adjectives differ from the infinitive and the present tense forms in lacking the thematic suffix *-nq-*. The resultative adjectives show the same thematic vowel as the past tense forms in (20d–e) and exhibit the same root allomorphy as the past tense form in (20f).

- (20) (a) wychudnie 'he/she/it will become thin' – wychudnąć 'to become thin' – wychudł 'he became thin' – wychudły 'emaciated, that has grown thin'
- (b) zdechnie 'it will die' – zdechnąć 'to die (of animals)' – zdechł 'it died' – zdechły 'dead'
- (c) zwiędnie 'it will wither' – zwiędnąć 'to wither' – zwiędł 'it withered' – zwiędły 'withered'

⁷ Occasionally one can note a slight difference between the past tense third person masculine singular form and the resultative adjective, as in (20c) or in the case of the verbs containing the root *chod-*. This is due to the regularizing force of the inflectional paradigm in the corresponding adjectives. For example, the root of the verbs *przychodzić* 'to come' and *odchodzić* 'to leave, to depart' exhibits the allomorph *-szed-* in the masculine singular past-tense forms *przyszedł* 'came-3sg.masc', *przyszedłem* 'came-1sg.masc', *odszedł* 'departed-3sg.masc' and *odszedłem* 'departed-1sg.masc', respectively. In the remaining past tense forms the root surfaces as *-sz-*, e.g., *przyszła* 'came-3sg.fem', *przyszłyście* 'came-3pl.fem' and *przyszliśmy* 'came-1pl.masc'. The same shape of the root occurs in the related resultative *-t-* adjective, as in *przyszły* 'future-nom.sg.masc' and *przyszła* 'future-nom.sg.fem'. The regularizing influence of the inflectional adjectival paradigm is also responsible for the difference between the past tense form *zwiędł* 'withered-3sg.masc' (exhibiting a different vowel than *zwiędła* 'withered-3sg.fem') and the resultative adjectives *zwiędły* 'withered, faded'.

- (d) *dojrzeje* 'he/she/it will ripen, mature' – *dojrzeć* 'to ripen, to mature' – *dojrzał* 'he ripened, grew mature' – *dojrzały* 'ripe, mature'
- (e) *posiwieje* 'he/she/it will grow grey' – *posiwieć* 'to grow grey' – *posiwał* 'he grew grey' – *posiwały* 'that has become grey'
- (f) *umrze* 'he/she/it will die' – *umrzeć* 'to die' – *umarł* 'he died' – *umarli* 'dead'

Similar examples of formal affinity can be observed between past tense forms and the so-called passive adjectives of possibility ending in the sequence *-aln(y)*, derived from imperfective verbs and discussed in Szymanek (1985). Both types of formations contain, for instance, the derived imperfective (DI) suffix *-ywa-* whereas the present tense forms exhibit *-uj-* as the allomorph of the suffix in question.

- (21) (a) *porównuje* 'he/she/it compares' – *porównywać* 'to compare' – *porównywał* 'he compared-impf.' – *porównywalny* 'comparable'
- (b) *przewiduje* 'he/she/it predicts' – *przewidywać* 'to predict' – *przewidywał* 'he predicted-impf.' – *przewidywalny* 'predictable'
- (c) *rozwiązuje* 'he/she/it solves' – *rozwiązywać* 'to solve' – *rozwiązywał* 'he solved-impf.' – *rozwiązywalny* 'solvable'

What is even more puzzling, the putative suffix *-aln-* can be found only in adjectives derived from imperfective verbs containing the DI suffix *-ywa-*. Imperfective verbs with the thematic vowel *-i-/-y-* or with no thematic vowel, derive adjectives with the adjectivizing (non-passive) suffix *-n-* (cf. 22a), or in the less productive suffix *-liw-* (in 22b). Alternatively, such verbs may have no related passive adjectives of possibility, as in (22c):

- (22) (a) *palić* 'to burn' – *palny* 'combustible'
gnić 'to rot' – *gnilny* 'decaying, related to the process of decaying'
przewozić 'to transport' – *przewoźny* 'transportable'
pić 'to drink' – *pitny* 'drinkable'
- (b) *łączyć* 'to join' – *łączliwy* 'joinable'
kurczyć 'to shrink' – *kurczliwy* 'shrinkable'
tłoczyć 'to stamp (metal)' – *tłoczliwy* 'drawable'
- (c) *liczyć* 'to count' – **liczalny* 'countable' (listed as 'archaic' in *SJP*)

Therefore, Szymanek (1985, 116) analyses *-al-* in *-aln-* as part of the past tense stem and formulates the following constraint on *-n-* suffixation:

- (23) Polish "adjectives of passive possibility" are productively derived from the third-degree (*-ł-*) stem of DI transitive verbs.

As illustrated in (24a–b), adjectives in *-aln-* can be occasionally related to verbs with the thematic vowel *-i/-y-* or *-e-*. The verb with the thematic vowel *-a-* in (24c) derives, in turn, an adjective in *-eln-*. Such exceptions are, however, very rare.

- (24) (a) policzyć (perf.) ‘to count’ – policzył ‘he counted up’ – policzalny ‘countable’
 (b) słyszeć ‘to hear’ – słyszał ‘he heard’ – słyszalny ‘audible’
 (c) czytać ‘to read’ – czytał ‘he read’ – czytelny ‘readable, obvious’

It was shown in (4) in section 2.1 and is illustrated again below in (25) that *-t-* adjectives can occur with resultative adjuncts, directional PPs and past-time expressions such as *przed chwilą* ‘a moment ago’, *dziś rano* ‘this morning’ and *zimą* ‘in winter’, similarly to passive participles and present *-qc(y)* participles (in 25d–e). This is not possible with the other deverbal adjectives (those exhibiting the suffix *-(al)n-*, *-liw-*, *-cz-*, *-qc-*) illustrated in (26). The latter adjectives can be combined neither with temporal phrases that can modify perfective finite verbs (e.g., *w ciągu pięciu minut* ‘in the course of five minutes’) nor with temporal adverbials possible with imperfective verbs (e.g., *przez pięć minut* ‘for five minutes’).

- (25) (a) przybyły przed chwilą na zamek posłaniec ‘a messenger who arrived at the castle a moment ago’
 (b) zgniłe zimą ziemniaki ‘potatoes that went bad during the winter’
 (c) rozkwitłe dziś rano w ciągu godziny tulipany ‘tulips that opened this morning within an hour’
 (d) rozbita przed chwilą przez ciebie na drobne kawałki szyba ‘the window pane that was broken by you into small pieces a moment ago’
 (e) płaczący wczoraj chłopiec ‘the boy who was crying yesterday’
- (26) (a) ?*widzialna wczoraj tęcza ‘the rainbow that was visible yesterday’
 (b) ?*palny na popiół w ciągu pięciu sekund/ przez pięć sekund materiał ‘material that can burn and turn into ash within five seconds/ for five seconds’
 (c) *ożywcze w ciągu pięciu minut/przez pięć minut powietrze ‘(breath of) air refreshing within five minutes/for five minutes’
 (d) ?*kłótlive w ciągu godziny/przez godzinę sąsiadki ‘(female) neighbours quarrelsome within an hour/for an hour’
 (e) *interesujący wczoraj film ‘yesterday’s interesting film’

If both *-aln-* adjectives and *-t-* adjectives are derived from the past tense (the third degree) stems of related verbs, it may come as a surprise that only the latter derivatives incorporate some notion of tense and time. An explanation for this difference can be provided once it is assumed that *-aln-* forms derive from the third degree (past tense) stems while *-t-* adjectives are derived by means of conversion from past (perfect) participles. It can be postulated that conversion of verbal past participles into deverbal resultative adjectives involves inheritance of event structure. The concept of event structure will be explicated in section 5. First I will mention briefly some problems involved with the proposal of adjectivization of past participles.

4. Conversion of past participles

4.1. Diachronic conversion

The analysis of *-t-* resultative adjectives as adjectivized active past participles is common in diachronic studies of Polish,⁸ for instance in Bartnicka (1970) or Długosz-Kurczabowa–Dubisz (1998). The verb in proto-Slavonic language had two stems: the present tense stem and the infinitive stem. There were three types of past participles, all of them formed from the infinitive stem: two active past participles (the so-called second past participle terminating in *-lŭ*) and one passive past participle (terminating in *-nŭ*). The second active past participle acted as a constituent in various periphrastic inflectional forms. It was combined with an appropriate auxiliary verb in complex past tense forms, pluperfect forms, complex future tense forms and conditional mood forms. In Old Polish analytical past tense forms were restructured: auxiliary verbs fused with active past participles, giving rise to simple (mono-lexemic) past-tense forms, e.g., *osiadł jeś* → *osiadłeś* 'you-sg. settled down', *padł jeśm* → *padłeśm* (*padłem*) 'I-fell (down)', *przyszli jesmy* → *przyszliśmy* 'we-came'. The third person singular auxiliary verb *jest* was lost in past tense forms and, as a result, the respective past tense form was identical to the second active past participle, e.g., *uciekał jest* → *uciekał* 'he-ran'.

⁸ As pointed out to me by Christina Y. Bethin (p.c.), a different direction of change was observed much earlier. The Proto-Indo-European resultative participle became the source for passive participles in Slavic (see also Haspelmath 1994, 161).

4.2. Synchronic adjectivization

A potential counterargument to the synchronic analysis of *-t-* adjectives as derived through conversion from *-t-* participles is provided by the fact that—according to many morphologists (including Laskowski 1984)—active past participles no longer belong to the inflectional paradigm of the Polish verb. As shown in (18) in the preceding section, Laskowski (1984) assumes that past tense forms, conditional forms and periphrastic future tense forms of a given Polish verb share a common element because they are all based on the third degree verb stem. However, the data in (18) can be analyzed in a different way. It can be claimed, as in Tokarski (1973), that the inflectional forms in question contain the past (perfect) participle, which combines either with the future auxiliary *być* ‘to be’ or with clitic-like conditional morpheme and person endings. If *widziała-* in *widziałas* ‘you-saw-sg.fem’ and in *widziałabyś* ‘you-would see-sg.fem’ is treated as an independent component of compound forms, it is to be predicted that person endings and conditional endings can be separated from past participles and that they are able to attach to other phonological hosts in (27).

- (27) (a) *żebyś widziała* ‘so that you would see’ (*że* ‘that’+conditional particle+ person ending) (Past Participle)
- (b) *tyś widziała* ‘you saw’ (*ty* ‘you-sg’+person ending) (Past Participle)
- (c) *będziesz widziała* ‘you-sg.fem will see’ (*be-fut.*+person ending) (Past Participle)

Such an analysis provides a neat explanation for the split between the marking of the categories of number, gender and person in finite verb forms. While number and gender are uniformly marked on the past participle forms by inseparable (immovable) inflectional affixes, person markers can either be affixed to auxiliary verbs or attached as clitics onto some likely phonological hosts (such as conditional particles, emphatic particles, adverbs or personal pronouns). Thus, the recognition of non-passive past participles as synchronic verb forms leads to a more satisfactory description of the peculiarities of Polish verb conjugation. Consequently, past participles recognized as synchronic entities can function as bases for lexical derivation.

I assume, as in Cetnarowska (1998), that conversion is an affixless morphological operation which results in a change of semantic and/or syntactic characteristics of the derivational base. If the Separation Hypothesis is adopted, as advocated in Szymanek (1985) and Beard (1995), conversion can be treated as L-derivation (semantico-syntactic change) without any concomi-

tant M-derivation (morphophonological change, such as affixation). The syntactic effect involved in the derivation of *-t-* adjectives is the relabelling of the past participle as an adjective (V → Adj). It involves a modification of inflectional properties of the base. While verbal past participles are inflected for gender and number, non-passive resultative adjectives agree in gender, number and case with their head nouns. The semantic effect of adjectivization is the change from denoting events to denoting properties (from dynamic to stative interpretation).⁹

In spite of denoting properties, past participles converted into adjectives exhibit event structures, similarly to finite verbs. In section 5 below I will present some basic information about the notion of event structure.

5. Event structure

5.1. Preliminaries

As proposed in Grimshaw (1990), each verb has an event structure associated with it (as a part of its lexical entry). The event structure specifies the aspectual category of the verb and determines the availability of adjuncts and temporal adverbials of a particular type with the verb.

Vendler (1967) identified four types of eventualities¹⁰ and, consequently, four classes of verbs denoting them: states (lying, loving), activities (running, walking), accomplishments (e.g., building a house) and achievements (dying, arriving, winning a contest). Achievements are momentaneous (non-durative) while accomplishments denote events of some duration. Both achievements and accomplishments imply a change of state or location of the participant(s) in the event. They are telic events, in other words they imply the existence of

⁹ In Cetnarowska (1998) I analyze passive resultative adjectives as derived through conversion from verbal passive participles, as is argued for English in Levin – Rappaport (1986). This instance of conversion involves a modification of the meaning of the base (from dynamic to stative) but results in no change of the inflectional paradigm of the base. A different way of deriving adjectival passives is adopted for instance in Marantz (1999), where both eventive (verbal) and stative (adjectival) passive are formed syntactically. Eventive passives are derived when the participle head (Pass) attaches above little *v* while stative (adjectival) passives involve the head Pass attaching below little *v* (at the level of root).

¹⁰ Vendler (1967) talks about types of events. However, in the later literature on the subject the term ‘event’ is often restricted to dynamic events. The term ‘eventuality’ is then employed to encompass both states and dynamic events.

a terminus (inherent end-point). Activities (processes) are durative but they do not denote any change or do not imply the existence of an inherent end-point. The occurrence of temporal adverbials in a clause is indicative of the aspectual properties of verbs (or verb phrases). Activities can be modified by durative phrases such as *for a year*. They do not allow the phrase *in a year* (as in *John studied physics for a year/*in a year*). Accomplishments are not modified by durative temporal adverbials but occur easily with *in three months* phrases, e.g., *They built a house in three months/*for three months*. Achievements are felicitous with punctual temporal adverbials, such as *at ten o'clock* though they also allow *in an hour* type of adverbials, as in *The bomb exploded at ten o'clock/The bomb will explode within an hour*.

Aspectual properties of verbs reflected in Vendler's typology are referred to as the Aktionsart of verbs in Verkuyl (1993) (though the term "Aktionsart" has a different usage in the linguistic literature on Slavic aspect, see Rothstein 1970 and Comrie 1976).¹¹ Verkuyl distinguishes between the "inner" and "outer" aspect of a predicate. The "inner" aspect is determined by the verb's aspectual (Aktionsart) properties, the properties of the verb's arguments and the presence of directional prepositional phrases. The "outer"—or "presentational"—aspect of a clause in English is computed on the basis of the "inner" aspect and the aspectual properties of tense and temporal adverbials.

In Slavic languages, such as Polish, the "outer" aspect is determined mainly by the aspectual properties of verbs, i.e., their inherent telicity reflected in the grammaticalized opposition between perfective and imperfective verb forms (see Schoorlemmer 1995). Atelic events are normally associated with imperfective verb forms while telic events are denoted by perfective verbs.¹²

¹¹ The term "Aktionsart" in the Slavic literature refers to various modes of action such as, for instance, 'limitative' (*popisać* 'to write for a while'), 'perdurative' (*przesiedzieć* 'to sit up (e.g., all night)') or 'saturative' (*najeść się* 'to eat until one feels full'). They are expressed by non-grammaticalized prefixes, in contrast to the aspectual distinction between the perfective and imperfective verb forms expressed by grammatical suffixes or prefixes.

¹² The relationship between telicity of a clause and the grammatical aspect of verbs is not, however, exceptionless. Imperfective verbs can occasionally denote telic events, as in the sentence *Kto malował ten portret?* 'Who painted-impf. this portrait?' (see Schoorlemmer 1995, 112ff).

5.2. Aspectual classification of bases for *-ł(y)* derivation

The fact easily noticeable about non-passive resultative *-ł(y)* adjectives is that they are derived from perfective verbs only.¹³ Consequently, the perfective verb *przybyć* 'to arrive-pf' has a related adjective *przybyły* 'arrived', while the derived imperfective verb *przybywać* 'to arrive-impf.' has no corresponding adjective **przybywały* 'that has been arriving, that used to arrive'. However, not every past participle of a perfective verb can undergo conversion into adjectives. There are no resultative adjectives **krzyknęły* 'that has shouted', **błysnęły* 'that has flashed', **pobiegały* 'that has run about' or **poczekwały* 'that has waited' related to the perfective verbs *krzyknąć* 'to shout', *błysnąć* 'to give a flash', *pobiegać* 'to run about' and *poczekać* 'to wait (for a while)'. The ill-formedness of these adjectives can be attributed to the aspectual properties of related verbs.¹⁴ These verbs cannot occur with the temporal modifiers such as *w trzy godziny* 'within three hours' but are compatible with the phrases *przez trzy godziny* 'for three hours' or *o trzeciej* 'at three o'clock', as shown by the sentences *Światło błysnęło *w trzy godziny/*przez trzy godziny/kwadrans po trzeciej* 'The light flashed in three hours/for three hours/at a quarter past three' and *Dzieci pobiegały *w trzy godziny/przez trzy godziny/*kwadrans po trzeciej* 'The children ran about in three hours/for three hours/at a quarter past three'. Such diagnostic sentences imply that the verbs in question are atelic. Consequently, they are unable to derive resultative adjectives, in spite of being perfective verb forms.

As far as the characterization of verbal bases for Polish *-ł-* adjectives in Vendler's typology is concerned, many of those verbs denote "accomplishments", e.g., *dojrzeć* 'to ripen', *wychudnąć* 'to grow thin', *spróchnieć* 'to rot', *owrzodzieć* 'to become affected with ulcers'. This is shown in (28) for the verb *dojrzeć* 'to ripen', where it occurs with the adverbial *w ciągu paru dni* 'within

¹³ A handful of exceptions was listed in section 2.3 in (12) and (14). However, derivatives from imperfective and nonprefixed verbs typically exhibit semantic noncompositionality.

¹⁴ An additional morphological restriction on the formation of *-ł-* adjectives is the requirement for prefixed bases. Observe that the non-prefixed telic verbs *pęknąć* 'to burst' and *minąć* 'to pass by, be over' have no related *-ł-* adjectives (while there are related *-n-/-t-* adjectives *pęknięty* 'burst open, cracked' and *miniony* 'past, bygone').

a couple of days' but is not fully felicitous with the durative phrase *przez cały tydzień* 'for a whole week'.¹⁵

- (28) (a) Jabłka w twoim sadzie dojrzeją w ciągu paru dni
 apples in your orchard ripen-fut.3pl in course several days
 '(The) apples in your orchard will ripen within several days.'
- (b) (*)Jabłka w twoim sadzie dojrzeją przez (cały) tydzień
 apples in your orchard ripen-fut.3pl for whole week
 '(The) apples in your orchard will ripen for the whole week.'

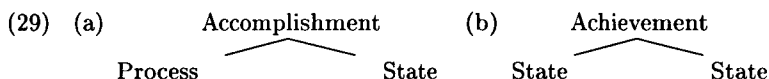
Apart from accomplishments proper, verbs related to *-ł-* adjectives include those denoting "degree achievements" (cf. Dowty 1979), such as *posiwieć* 'to grow greyish', *pociemnieć* 'to become darker'. They denote a change in a particular direction but, instead of denoting the achievement of an inherent terminus, they imply reaching one of intermediate "end-points". For instance, the verb *pożółknąć* 'to become slightly yellow' differs from the related verb *zżółknąć* 'to become yellow' since the former verb denotes some degree of change in colour of a given object and the latter denotes a complete change of colour, i.e., becoming yellow. Both of those verbs form resultative *-ł-* adjectives, i.e., *pożółkły* 'that has become slightly yellow' and *zżółkły* 'that has grown (completely) yellow'.

Furthermore, there are resultative adjectives related to "achievement verbs", i.e., predicates denoting a change of state (or location) that is instantaneous, e.g., *owdowieć* 'to become a widow/widower', *umrzeć* 'to die', *upaść/paść* 'to fall down' or *ochrypnąć* 'to become hoarse'.

Within the theory of event structure put forward in Pustejovsky (1991), non-stative eventualities exhibit internal structure. Achievements and accomplishments are referred to as transitions and are analysed as consisting of two subevents. Painting a portrait, for instance, involves an activity (process in Pustejovsky's terminology) of painting as the first subevent and the resulting

¹⁵ The reason why (28b) is preceded with a star in parenthesis is that temporal adverbials in Polish can change their interpretation under the influence of the perfective or imperfective (telic/atelic) status of the accompanying finite verb. In colloquial Polish the sentence *Przez tydzień jabłka w naszym sadzie dojrzeją.* is perfectly acceptable since the phrase *przez tydzień* can be interpreted as meaning 'within a week' (in addition to the expected 'for a week' reading). There is a difference, thus, between English and Slavic languages (such as Polish). In English temporal adverbials disambiguate between telic and atelic predicates (*He played sonatas for an hour/He played the sonata in an hour*). In Polish the telicity of the predicate is signalled by the verb morphology and, in order to avoid the ill-formedness of the sentence due to the choice of the inappropriate aspectual temporal adverbial, the interpretation of the adverbial may be modified.

state (i.e., the portrait being finished) as the second subevent. An achievement, e.g., the predicate *dying*, involves two subevents: the initial state (of being alive) and the resulting state (of being dead).



Predicates can give rise to resultative adjectives if they involve the resulting state subevent in their event structure.

5.3. Event structure of adjectives

Sleeman–Verheugd (1998) argue that premodifying participles/deverbal adjectives in French are derived in the lexicon but, like related verbs, they can have both argument structure and/or event structure.

A similar proposal can be put forward for Polish. Though deverbal *-t(y)* formations, by virtue of being adjectives, denote a property (state) of the head nouns they modify, they imply that some previous event took place. This implication is “translated into” the event structure grid present in their lexical entries. The occurrence of resultative phrases and directional PPs follows then from (is licensed by) the aspectual properties of *-t(y)* adjectives reflected in their event structure. Since they have a two-part event structure, they are compatible with phrases describing the end-point (i.e., the end state or the end location) and with temporal adverbials that can accompany telic events.

Having provided a plausible explanation for why adverbials and resultative adjuncts are **possible** with deverbal adjectives, it then remains to account for why modifiers and adjuncts are **obligatory** with certain deverbal adjectives and participles. Following Ackerman–Goldberg (1996), I use “#” to indicate phrases unacceptable in a “neutral” (i.e., non-contrastive) context.

- (30) (a) #rozgorzały spór ‘the quarrel that has broken out’
 (b) rozgorzały na nowo spór ‘the quarrel that has broken out again/anew’
 (c) #wylęęte ptaki ‘the hatched birds’
 (d) nowo wylęęte ptaki ‘newly-hatched birds’
 (e) #odosnęę gałęęie ‘the twigs that have grown out’
 (f) odosnęę na wiosnęę gałęęie ‘the twigs that have grown out in spring’

- (g) #powstały uniwersytet 'the university which has been created'
 (h) nowo-powstały uniwersytet 'the university which has been created recently'
 (i) #zbudowany szpital 'the built hospital'
 (j) zbudowany na miejscu wysypiska śmieci szpital 'the hospital built on the former dump-site'

Grimshaw–Vikner (1993) discuss a similar phenomenon in English, exemplified for adjectival passive participles in (31) below.

- (31) (a) #a created house
 (b) a carefully created house
 (c) #a built house
 (d) a recently built house
 (e) #a recorded message
 (f) a badly recorded message

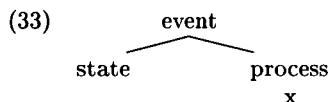
Grimshaw and Vikner put forward the hypothesis that there is a difference in the role of the head noun (i.e., the DP bearing the role of Theme/Patient) in the event structure of verbs of creation (such as *create*) and verbs denoting other types of transitions (e.g., *record*). The *y* variable denoting the object participant is present in (“identifies”) both subevents of the verb *record*, i.e., the process (activity) and the resulting change of state. In the case of the verb *create*, in contrast, *y* “identifies” only the second subevent since it does not come into being previously to the completion of the event.

- (32) (a) x recorded y (b) x created y
- | | |
|--|--|
| | |
|--|--|

According to Grimshaw and Vikner, both subevents of an accomplishment must be “identified” by some element in a clause. Since the head noun in a phrase such as *a created house* in (31a) identifies only the second (resulting state) subevent, the phrase is infelicitous in the absence of some temporal or

manner adverbial which would serve to “identify” the first subevent (i.e., the process).¹⁶

Grimshaw and Vikner’s proposal can account for the infelicity of unmodified pronominal passive resultative adjectives in Polish, such as #*zbudowany szpital* ‘built hospital’ in (30i). Their account can be extended to the cases of the infelicitous non-passive resultative adjectives in (30), e.g., #*rozgorzały spór* ‘the quarrel that has broken out’ in (30a). A probable event structure for the inchoative intransitive verb *rozgorzeć* ‘to break out’ is given in (33).¹⁷



The head noun *spór* ‘quarrel’ in (30a) “identifies” the second subevent of the intransitive verb *rozgorzeć* ‘to break out’. The first subevent requires “identification” by an adjunct element, e.g., *na nowo* ‘anew, again’.

In the next section I will consider the hypothesis that verbs which give rise to non-passive resultative adjectives, such as *rozgorzeć* ‘to break out’ or *powstać* ‘to arise’ are unaccusative verbs.

6. The Unaccusative Hypothesis and the conversion of participles

Additional evidence supporting the analysis of non-passive resultative *-t-* adjectives as derived by conversion from past participles in Polish can be provided

¹⁶ A different account for the infelicity of phrases such as *a built house* is offered in Ackerman – Goldberg (1996). They propose the Non-redundancy Constraint which reads: “If the referent of the head noun, N, implies a property P as part of its frame-semantic or encyclopedic knowledge, then an APP is not allowed to simply designate P; it must be further qualified.” (Ackerman – Goldberg 1996, 21).

¹⁷ Inchoative verbs denote the inception of a process. I follow van Hout (1996, 93) in analyzing their event structure as consisting of a state as the initial subevent and a process as the final subevent, e.g., the state of not being on fire and the process of burning in *rozgorzeć* ‘to start to burn; to break out’. In the case of *rozgorzały na nowo spór* ‘the quarrel that has broken out anew’ in (30b), the adjunct *na nowo* ‘anew’ provides information about the whole transitional event but this is not crucial for the “identification” of the second subevent. The term “inchoative” is used in a different sense in, among others, Levin – Rappaport Hovav (1995), who regard alternating verbs, such as *break* or *open* in English, as exhibiting a causative and inchoative usage. The structure in (33) implies that resultative adjectives can be derived from verbs denoting eventualities whose second subevent is a resulting process rather than a state.

from cross-linguistic studies. It is suggested for Dutch (Hoekstra 1984) that the past participles of unaccusative verbs (in contrast to unergatives) can undergo conversion into adjectives and occur as premodifiers, as in (34):

- (34) (a) de gevallen man (unaccusative verb)
 'the fallen man'
 (b) *de gewerkte man (unergative verb)
 'the worked man'

A conversion analysis for the English adjectival participles in (35) is espoused in, among others, Levin–Rappaport (1986). As shown in (35), it is only the past participles of unaccusative verbs in English which can occur prenominaly.

- (35) (a) the deceased man (unaccusative verb)
 (b) the wilted flowers (unaccusative verb)
 (c) *the run man (unergative verb)
 (d) *a coughed patient (unergative verb)

Markantonatou (1995) regards conversion of the perfect (past) participles into adjectives as one of the diagnostics of the unaccusative character of related verbs in Greek, e.g., *peθamenos* 'dead' from *peθeno* 'I die'.¹⁸ If I can show that non-passive resultative *-t-* adjectives in Polish are related to unaccusative verbs, then the proposal of deriving them from the corresponding past participle forms through conversion would fit the pattern observed cross-linguistically.

Since an in-depth discussion of unaccusativity is beyond the scope of the present paper, I will present only the most basic information about the Unaccusativity Hypothesis in section 6.1 below. In section 6.2 I will show that the semantic classes of verbs in Polish which have related *-t-* adjectives are good candidates for unaccusatives. In section 6.3 I will point out that the analysis of the bases for *-t-* adjectives in terms of their event structure further supports their treatment as unaccusatives.

¹⁸ Haspelmath (1994, 157ff) provides examples from Mongolian, Kanuri, Margi, Hungarian, Turkish, Arabic, Mam and Panare to show that participles/deverbal adjectives can modify subjects of unaccusative verbs.

6.1. The intransitivity split

The Unaccusativity Hypothesis, put forward originally in Perlmutter (1978) and elaborated upon in Burzio (1986), predicts the universal (cross-linguistic) existence of two classes of intransitive predicates: unergatives and unaccusatives. They differ in their syntactic behaviour, for example in the selection of the auxiliary verb in perfect tenses and the ability to form impersonal passives. In Dutch, unaccusative verbs require the auxiliary verb *zijn* 'to be' in perfect tenses and do not undergo impersonal passivization. In contrast, unergative and transitive verbs can occur in the impersonal passive construction and select the auxiliary *hebben* 'to have' in perfect tenses. Subjects of unaccusative verbs pattern syntactically with objects of transitive verbs. In Italian, for instance, the clitic *ne* 'of them' can be extracted out of a DP which functions as a direct object of a transitive verb or a postverbal subject of an unaccusative verb (Burzio 1986). Therefore, a common assumption made in the literature on unaccusativity is that subjects of unaccusative verbs are internal arguments (D-structure objects) while subjects of unergative verbs and transitive verbs are external arguments (cf. Burzio 1986; Tenny 1994). Another line of research on the intransitivity split emphasizes the semantic difference between unaccusative and unergative verbs (Zaenen 1993; Levin–Rappaport Hovav 1995), which results in their distinct syntactic behaviour. Unaccusative verbs denote a change of state and/or location and their single argument bears the thematic role of Patient or Theme. Subjects of unergative verbs are typically Agents. Consequently, in Hale–Keyser (1993) unaccusative verbs are analyzed as lacking an agentive (vP) projection above VP. I will show in section 6.2 below that the semantic characterization of verbs which form non-passive resultative *-t-* adjectives in Polish indicates the possibility of analyzing them as unaccusatives. Note that in Polish the class of intransitive verbs exhibiting the unergative/unaccusative split includes both “intransitives proper”, such as *upaść* ‘to fall’, as well as verbs occurring with the reflexive clitic *się*, e.g., *roztopić się* ‘to melt’.

6.2. Semantic characterization of bases for *-t-* adjectives

There is a semantic restriction on verbal bases which give rise to resultative *-t-* adjectives in Polish. The “surface” subjects of such verbs denote Patients and Themes rather than Agents. As illustrated in (36), verbal bases for *-t-* derivation denote a change of state or a change of location. In this respect Polish

resultative adjectives resemble attributive perfect participles in German (as discussed in, among others, Wunderlich 1997) and adjectival past participles in English (cf. Levin 1993).

(36) Verbs which have related non-passive resultative *-ł-* adjectives

- I. Verbs denoting a change of state: *omdleć/zemdleć* 'to faint', *opuchnąć/spuchnąć* 'to become swollen', *oślepnąć* 'to become blind', *pozółknąć* 'to become slightly yellow', *sczernieć* 'to grow black', *stępieć* 'to become dull, less sharp', *schudnąć/wychudnąć* 'to grow thin', *zachrypnąć* 'to grow hoarse', *zmatowieć* 'to become dull and opaque', *znieuchomić* 'to become motionless'.
- II. Verbs of inherently directed motion: *przybyć* 'arrive', *upaść* 'fall', *zbiec* 'to escape'.
- III. Verbs of disappearance: *poleć, paść* 'to die in a battle', *przepaść* 'to disappear', *umrzeć, zgasnąć* (poet.) 'to die', *wygasnąć* 'to expire'.
- IV. Verbs of appearance and occurrence: *powstać* 'to come into being', *wyniknąć* 'to ensue, to result', *zaistnieć* 'to come into being'.
- V. Verbs of assuming position: *obwisnąć* 'to hang down', *przyłec* 'to come to adjoin'.

The verb classes in (36) constitute a subset of predicates which are generally predicted to be unaccusative cross-linguistically on the basis of their meaning (as proposed in Perlmutter 1978).

Similar classes of verbs exhibit the behaviour of unaccusative predicates in English (see Levin–Rappaport Hovav 1995). In Cetnarowska (2000) I show that the verbs listed in (36) do not occur in the impersonal *-no/-to* construction in Polish, which also suggests their analysis as unaccusatives.¹⁹

6.3. Unaccusativity and event structure

It has been proposed in the literature that unaccusativity is based on aspectual distinctions. Van der Putten (1997, 120) states that all unaccusative verbs are telic, i.e., they are either accomplishments or achievements. Verbs which give rise to resultative *-ł-* adjectives in Polish are telic (are shown to be achieve-

¹⁹ There are some differences between the set of verbs which form *-ł-* adjectives and the set of verbs impossible in the *-no/-to* impersonal sentences since each of the construction exhibits additional restrictions (cf. Cetnarowska 2000).

ments or accomplishments in section 5.2).²⁰ They exhibit the internal subevent structure of transitional (telic) events, as shown in (29), (32) or (33). Their aspectual properties are compatible with their analysis as unaccusatives.

Tenny (1994) puts forward the hypothesis that only internal arguments can “measure out” events, where “measuring out” implies that the argument marks the temporal end-point of the event. *His milk* in the sentence *Thomas drank up his milk* “measures out” (delimits) the event described by the verb since the end of the event corresponds to the state of the milk being consumed. Tenny argues that if the surface subject delimits (measures out) the event, the verb is unaccusative. In the case of all the verbs from which non-passive resultative *-ł-* adjectives are formed, the surface subject acts as an argument measuring out the event over time. This can be shown for the verbs *zgnić* ‘to become rotten’ and *zamarznąć* ‘to freeze’ and , which derive the adjectives *zgniły* ‘rotten’ and *zamarzły* (or *zamarznięty*) ‘frozen’. In (37) I apply the equivalents of the adverbial tests used for English in Tenny (1994, 63) to show the existence of a measuring argument.

- (37) (a) Jezioro w połowie/częściowo zamarzło.
‘The lake froze halfway/partly.’
- (b) Jezioro zamarzło w ciągu tygodnia.
‘The lake froze in a week.’
- (c) Jabłko w połowie/częściowo zgniło.
‘The apple became halfway/partly rotten.’
- (d) Jabłko zgniło w ciągu trzech dni.
‘The apple became rotten within three days.’

The surface subjects of the verbs in (37) measure change along a temporal scale and mark its terminus, hence the felicity of the adverbial element *w połowie* ‘halfway’ and of the temporal phrases *w ciągu miesiąca/w ciągu trzech dni* ‘within a month/within three days’.

If it is only unaccusative verbs in Polish which give rise to resultative *-ł-* adjectives, the adoption of the conversion analysis for the derivation of

²⁰ There are a few atelic verbs of state which have related *-ł-* adjectives, e.g., *trwać* ‘to last, to endure’ giving rise to *trwały* ‘permanent, constant’. Such adjectives tend to exhibit semantic idiosyncrasies, as was shown in section 2.3. However, it is crucial to note that cognate verbs in Russian have been analyzed as unaccusatives, which suggests that the link between the unaccusativity of a verb and its ability to form the non-passive resultative *-ł-* adjective can override the requirement for the telicity of the base.

those adjectives becomes more plausible since it follows the pattern observed cross-linguistically.

7. Conclusion

In the preceding paper I have argued that non-passive resultative adjectives in Polish, such as *zwiędły* 'faded', should be derived through conversion (adjectivization) of past participles. I demonstrated (in section 2) that the syntactic derivation of non-passive resultative adjectives is implausible in view of the occasional unpredictability of the morphological shape of the derivative (i.e., the occurrence of the *-t-* or *-n-/-t-* suffix) and the semantic noncompositionality of numerous *-t-* and *-n-/-t-* adjectives. I showed (in section 3) that resultative *-t-* adjectives frequently exhibit stem allomorphy, which further supports the idea of their lexical (i.e., non-syntactic) derivation. I argued (following Tokarski 1973) that past participles should be recognized as belonging to the inflectional verb paradigm, although they occur only in analytic tense or mood forms. I illustrated the contrast between adjectives of passive possibility (such as *palny* 'combustible' and *powtarzalny* 'that can be repeated') and non-passive resultative adjectives (*zwiędły* 'withered' and *zbiegły* 'escaped'). While both types of adjectives show the same stem allomorphy as past tense forms of related verbs, it is only the latter type of adjectives which show the aspectual properties of related verbs. They are compatible with temporal phrases and resultative adjuncts of the same type as related verbs. I suggested that the lexical derivation of resultative *-t-* adjectives through conversion from past participles allows for the occurrence of the event structure with the adjectives in question. The event structure licenses the occurrence of temporal phrases such as *w ciągu trzech dni* 'within three days', of manner adverbials and resultative adjuncts. Furthermore, I pointed out that the infelicity of certain non-passive resultative adjectives in the absence of adverbial modification or resultative adjuncts can be accounted for in terms of their event structure (as was proposed for English deverbial adjectives/participles in Grimshaw-Vikner 1993).

Finally, I employed cross-linguistic evidence to make the conversion hypothesis more plausible. I proposed that intransitive verbs in Polish exhibit the split into unaccusative and unergative predicates. The semantic interpretation and the aspectual properties of the verbs which give rise to non-passive resultative adjectives suggest the unaccusative status of such verbs. Further research into the validity of the Unaccusativity Hypothesis in Polish is certainly required, although the data discussed in Cetnarowska (2000) indicate

that verbs from which resultative *-t-* adjectives can be derived exhibit other syntactic properties of unaccusatives.

References

- Ackermann, F. – Goldberg, A.E. 1996. Constraints on adjectival past participles. In: Goldberg, A.E. (ed.): *Conceptual structure, discourse and language*, 17–30. CSLI Publications, Stanford/California.
- Anderson, S.R. 1982. Where's morphology? In: *Linguistic Inquiry* 13: 571–612.
- Aronoff, M. 1994. *Morphology by itself*. MIT Press, Cambridge MA.
- Babby, L.H. 1973. The deep structure of adjectives and participles in Russian. In: *Language* 49: 349–60.
- Babby, L.H. 1998. Voice and diathesis in Slavic. Position paper for Comparative Slavic Morphosyntax, Bloomington, Indiana 5–7 June 1998.
- Bartnicka, B. 1970. *Adietywizacja imiesłowów w języku polskim* [Adjectivization of participles in Polish]. PWN, Warszawa.
- Beard, R. 1995. *Lexeme-morpheme base morphology: a general theory of inflection and word formation*. State University of New York Press, Albany.
- Burzio, L. 1986. *Italian syntax: a Government-Binding approach*. Reidel, Dordrecht.
- Cetnarowska, B. 1998. On inherent inflection feeding derivation in Polish. Ms.
- Cetnarowska, B. 2000. On the unergative/unaccusative distinction in Polish and the derivation of resultative adjectives. In: King, T.H. – Sekerina, I.A. (eds): *Formal approaches to Slavic linguistics. The Philadelphia Meeting 1999*, 78–96. Michigan Slavic Publications, Ann Arbor.
- Comrie, B. 1976. *Aspect*. Cambridge University Press, Cambridge.
- Długosz-Kuczabowa, K. – Dubisz, S. 1998. *Gramatyka historyczna języka polskiego* [A historical grammar of Polish]. Wydawnictwa Uniwersytetu Warszawskiego, Warszawa.
- Doroszewski, W. 1958–1969. *Słownik języka polskiego* [A dictionary of Polish language]. Volume 1–11. PWN, Warszawa.
- Dowty, D.R. 1979. *Word meaning and Montague grammar*. Reidel, Dordrecht.
- Drijkoningen, F. 1992. Derivation in syntax. In: Aronoff, M. (ed.): *Morphology now*, 48–68. State University of New York, New York.
- Grimshaw, J. 1990. *Argument structure*. The MIT Press, Cambridge MA.
- Grimshaw, J. – Vikner, S. 1993. Obligatory adjuncts and the structure of events. In: Reuland, E. – Abraham, W. (eds): *Knowledge and language II: lexical and conceptual structure*, 143–55. Kluwer, Dordrecht.

- Grzegorzczkova, R. 1979. Zarys słowotwórstwa polskiego. Słowotwórstwo opisowe. [A sketch of Polish word-formation. Descriptive word-formation]. PWN, Warszawa.
- Hale, K. – Keyser, S.J. 1993. On argument structure and the lexical expression of syntactic relations. In: Hale, K. – Keyser, S.J. (eds): *The view from building 20. Essays in linguistics in honor of Sylvain Bromberger*, 53–110. MIT Press, Cambridge MA.
- Haspelmath, M. 1994. Passive participles across languages. In: Fox, B. – Hopper, P.J. (eds): *Voice: form and function*, 151–77. Benjamins, Amsterdam.
- Hoekstra, T. 1984. *Transitivity*. Foris, Dordrecht.
- Hout, A. van 1996. Event semantics of verb frame alternations. A case study of Dutch and its acquisition. Ph.D. dissertation, University of Tilburg, Tilburg.
- Kallas, K. 1984. Słowotwórstwo przymiotników [Adjectival word-formation]. In: Grzegorzczkova, R. – Laskowski, R. – Wróbel, H. (eds): *Gramatyka współczesnego języka polskiego. Morfologia [A grammar of present-day Polish. Morphology]*, 408–55. PWN, Warszawa.
- Laskowski, R. 1975. *Studia nad morfonologią współczesnego języka polskiego [Studies on the morphophonology of present-day Polish]*. Ossolineum, Wrocław.
- Laskowski, R. 1984. Paradygmatyka [Inflection]. In: Grzegorzczkova, R. – Laskowski, R. – Wróbel, H. (eds): *Gramatyka współczesnego języka polskiego. Morfologia [A grammar of present-day Polish. Morphology]*, 171–303. PWN, Warszawa.
- Levin, B. 1993. *English verb classes and alternations: a preliminary investigation*. University of Chicago Press, Chicago.
- Levin, B. – Rappaport, M. 1986. The formation of adjectival passives. In: *Linguistic Inquiry* 17: 623–61.
- Levin, B. – Rappaport Hovav, M. 1995. *Unaccusativity. At the syntax-lexical semantics interface*. MIT Press, Cambridge MA.
- Marantz, A. 1999. Derivational morphology in an anti-lexicalist grammar. Talk given at Maryland Morphology Mayfest, College Park May 1999.
- Markantonatou, S. 1995. Modern Greek deverbal nominals: an LMT approach. In: *Journal of Linguistics* 31: 267–99.
- Nedjalkov, V.P. – Jaxontov, S.J. 1988. The typology of resultative constructions. In: Nedjalkov, V.P. (ed.): *Typology of resultative constructions*, 3–62. John Benjamins, Amsterdam.
- Perlmutter, D. 1978. Impersonal passives and the Unaccusative Hypothesis. In: *Proceedings of BLS* 4: 157–89. University of California, Berkeley.
- Pustejovsky, J. 1991. The syntax of event structure. In: *Cognition* 41: 47–81.
- Putten, F. van der 1997. *Matter and mind in morphology. Syntactic and lexical deverbal morphology in Dutch*. Ph.D. Dissertation, University of Leiden, Leiden.
- Rothstein, R.A. 1970. *The Polish verb*. Ms. Harvard University.

- Sadler, L. – Spencer, A. – Zaretskaya, M. 1997. A morphomic account of a syncretism in Russian deverbal nominalizations. In: *Yearbook of Morphology*, 180–215. Kluwer, Dordrecht.
- Scalise, S. 1988. Inflection and derivation. In: *Linguistics* 26: 561–82.
- Schoorlemmer, M. 1995. Participial passive and aspect in Russian. Ph.D. Dissertation, OTS, Utrecht University, Utrecht.
- SJP* = Doroszewski, W. 1958–1969. *Słownik języka polskiego* [A dictionary of Polish language]. Volume 1–11. PWN, Warszawa.
- Sleeman, P. – Verheugd, E. 1998. How reduced are reduced relatives? In: Kager, R. – Bezooijen, R. van (eds): *Linguistics in the Netherlands* 15, 187–9.
- Spencer, A. 1991. *Morphological theory*. Blackwell, Cambridge MA and Oxford.
- Szymanek, B. 1985. *English and Polish adjectives*. Redakcja Wydawnictw KUL, Lublin.
- Tenny, C.L. 1994. *Aspectual roles and the syntax-semantics interface*. Kluwer, Dordrecht.
- Tokarski, J. 1973. *Fleksja polska* [Polish inflection]. PWN, Warszawa.
- Vendler, Z. 1967. *Linguistics in philosophy*. Cornell University Press, Ithaca NY.
- Verkuyl, H. 1993. *A theory of aspectuality*. Cambridge University Press, Cambridge.
- Wunderlich, D. 1997. Participle, perfect and passive in German. In: *Working Papers of the SFB. Theorie des Lexicons*, vol. 99. Heinrich-Heine-Universität, Düsseldorf.
- Zaenen, A. 1993. Unaccusativity in Dutch: integrating syntax and lexical semantics. In: Pustejovsky, J. (ed.): *Semantics and the lexicon*, 129–61. Kluwer, Dordrecht.

Address of the author: Bożena Cetnarowska
Institute of English
University of Silesia
ul. Żytnia 10
41-205 Sosnowiec
Poland
cetnarow@uranos.cto.us.edu.pl

ASYMMETRIES:
CONSEQUENCES FOR MORPHOLOGICAL
CONFIGURATIONS AND PARADIGMS*

ANNA-MARIA DI SCIULLO

Abstract

In this paper we bring evidence to the effect that asymmetries are basic in the interpretation of the different sorts of objects generated by the grammar. We focus on the properties of morphological configurations, in particular the restrictions imposed on their derivations in order to show that they follow from our hypothesis. We discuss the properties of morphological paradigms, in particular the restrictions imposed on their actual members and their combination with roots in order to show that their properties are accounted for in terms of asymmetrical relations.

1. Local Asymmetry Hypothesis

Asymmetrical relations are basic in the grammar, as they contribute to determine restrictions on extraction, binding, and linear order (Chomsky 1981; 1995; 1998; Kayne 1984; 1994; Reinhart 1983; 1995; Hornstein 1995; Collins 1997). We propose a definition of asymmetrical relation that applies to areas not covered by current works on asymmetry and is compatible with the hypothesis that asymmetrical relations extend under the word level (Hale–Keyser 1993; Kayne 1994; Di Sciullo 1995; Keyser–Roeper 1995). We provide theoretical and empirical support to the hypothesis in (1), based on the definitions of ‘local domain’, given below.

(1) ASYMMETRY HYPOTHESIS

Grammatical relations are asymmetrical relations.

* A preliminary version of this paper was distributed at the 8th International Conference on Morphology, Budapest, 13-14/06/1998. I thank Noam Chomsky, Tom Roeper, Carol Tenny, Edwin Williams and two anonymous reviewers for discussions, comments and suggestions. This research is supported in part by the Social Sciences and Humanities Research Council of Canada, grant number 412-97-0016.

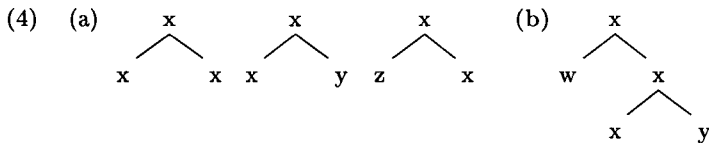
The Asymmetry Hypothesis bears on the way the performance systems, the Conceptual-Intentional (C-I) and the Acoustic-Perceptual (A-P) systems interpret the structural descriptions generated by the grammar, and thus the parts they are composed of. Even though this hypothesis has consequences through the grammar,¹ we focus here on morphological configurations and paradigms. In doing so, we bring evidence to the effect that asymmetries are basic in the interpretation of the different sorts of objects generated by the grammar. Let us start with the following definition.

(2) ASYMMETRICAL RELATION

r is asymmetrical =_{df} $(\forall x) (\forall y) (rxy \supset \sim ryx)$.

According to the definition in (2), an asymmetrical relation is a unidirectional relation r between two elements x and y such that rxy does not imply ryx .² Asymmetry applies to pairs of elements, such as (3), as well as to elementary and extended configurations, such as (4). This is not the case for sisterhood, which applies only to the elements in configurations such as (4a) and (4b), as well as asymmetrical c-command, which apply only to extended configurations such as (4b).

(3) (x, y)



The definition in (2) does not apply to arrays of lexical items, in particular, it does not apply to **numerations**, in the sense of Chomsky (1998). It is restricted

¹ For example, it leads to reinterpreting the predication relation, in terms of an asymmetrical relation; it also leads to reinterpreting the binding relations in terms of local asymmetrical relations between the members of chains. We will not discuss these issues here.

² The definition in (2) is distinct from other logical relations, such as symmetry and antisymmetry

(i) r is symmetrical =_{df} $(\forall x) (\forall y) (rxy \supset ryx)$

(ii) r is antisymmetrical =_{df} $(\forall x) (\forall y) ((rxy \wedge ryx) \supset x=y)$

These definitions are central in mathematics, as they distinguish the relation '=', which is a symmetrical relation, from '<', which is an antisymmetrical relation.

to pairs of grammatical elements in linguistic representations. Let us further define the notion of the local domain of an asymmetrical relation as follows.

(5) LOCAL DOMAIN

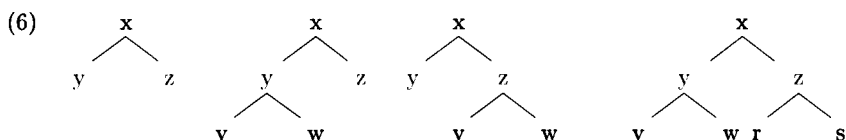
The local domain of $r(x, y)$ is the minimal domain where x and y are r -related.

According to (5), the local domain of an asymmetrical relation is not defined by a singular category, say by the presence of a given functional or lexical category, it is dependent on the configurations the grammatical elements are a part of.

In the following sections, we show that the notion of asymmetry covers a wide range of facts. In section 2, we focus on the properties of morphological configurations, in particular the restrictions imposed on their derivation. In section 3, we discuss the properties of morphological paradigms, in particular the restrictions imposed on their actual members and their combinations.

2. Morphological configurations

Morphological configurations are distinct from syntactic configurations (Di Sciullo–Williams 1987; Kayne 1994; Chomsky 1995; Di Sciullo 1996). They are not unorganized sets of elements; they consist of structured pairs of elements, as implemented in binary-branching trees, such as (6). They are reduced to head-adjunction structures at the interface with the performance systems, under a derivational approach, such as the one suggested in Kayne (1994), or under a representational approach, along the lines of Di Sciullo (1996) and Williams (1994).



Given the Asymmetry Hypothesis, we now take each element x, y, \dots in a morphological object m to be licensed by an asymmetrical relation r , such that:

- (7) For each x in m , there must be a y such that for (x, y) there is a r such that $r(x, y)$ is true in m .

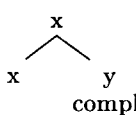
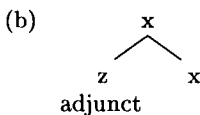
An asymmetrical relation unidirectionally relates types of grammatical entities to restrictors for these types. In the case of morphological configurations, the asymmetry holds for positions in binary branching structures. The relation

goes from types of positions, such as the head position, to restrictors, such as the complement and the adjunct positions. Let us further distinguish configurational r-relations from semantic r-relations, as in (8). For the latter relations, semantic types, such as predicate, are restricted by semantic restrictors, such as argument and modifier.

- (8) (a) $r_{configuration}$
 Configurational types include configurational restrictors.
- (b) $r_{semantic}$
 Semantic types include semantic restrictors.

Configurational r-relations are not co-extensive with semantic r-relations, as different semantic relations may be supported by the same configurational r-relation. This is the case for the predicate–argument relation which is supported by the head–complement configuration in phrasal structure and by the adjunct–head configuration in word-structure.

One theoretical consequence of our proposal is that it ensures autonomy as well as relatedness to configurational and semantic types. Another theoretical advantage of the notion of asymmetry is that it provides grammatical content to categories, such as complement and adjunct, which are nondistinct with respect to sisterhood, as illustrated below.

- (9) (a)  (b) 

- (10) (a) r_{compl}
 A head selects its complement.
- (b) $r_{adjunct}$
 An adjunct identifies its adjoiner.

As formulated above, the relation r_{compl} is asymmetrical as there is a unidirectional relation between two grammatical elements, a head and a complement, such that the head unidirectionally selects its complement. Likewise, the relation $r_{adjunct}$ is asymmetrical, as there is a unidirectional relation between two grammatical elements, an adjunct and an adjoiner, such that the adjunct identifies an unspecified feature of the category it is adjoined to.

2.1. Restrictions on the derivation of morphological configurations

One empirical consequence of the notion of asymmetry is that it contributes to derive the following restrictions on morphological configurations.

- (11) (a) The specifier (SU) cannot be the non-head.
 (b) The composition amongst non-heads is limited.

We consider the configurational properties of deverbal compounds, in order to derive the fact that the specifier (SU) cannot be licensed as the non-head in these configurations. We also discuss the restrictions on denominal verb formation in order to show that our hypothesis also extends to this class of morphological configurations.

2.1.1. Compounds

In productive compound structures, as it is more generally the case for morphological expressions, the $r_{adjunct}$ relation is canonical and overrides the r_{compl} relation, which may be licensed in the derivation. The $r_{adjunct}$ relation overrides the r_{compl} relation overtly in the structure of deverbal compounds of most languages with unemphatic SVO order.

- (12) (a) book-reading **English**
 novel-reader
- (b) kapn-o-kalierjia **Modern Greek**
 kapn-LV-kaliery- -i- -a
 'tobacco cultivate -ion nom. sg. fem'
- (c) bacc-e k^hojtii **Hindi**
 children-nom search-hab
 'children-searching'
- (d) knjig-o-nosa-∅ **Serbo-Croatian**
 book-LV-holder-nom
 'book-holder'
- (e) chleb-o-opiek-acz **Polish**
 bread-LV-toast-er/nom
 'toaster'

Interestingly, the $r_{adjunct}$ relation does not override the r_{compl} relation overtly in the structure of deverbal compounds in Romance and African languages, where the verb precedes its complement before Spell-Out.³

- (13) (a) schiaccia-noce **Italian**
 crack nut
 ‘nut-cracker’
- (b) ouvre-boîte **French**
 open box
 ‘can-opener’
- (c) awoko-busi **Yoruba**
 driver-bus
 ‘bus-driver’
- (d) okhie-elamhi **Yekhee**
 seller meat
 ‘meat-seller’

In both cases, the semantical $r_{predicate}$ relation holds, whether the internal argument (DO) of the verbal predicate is in complement position or in adjunct position at Spell-Out. In both cases, a complement (DO) can be in the non-head position. However, this is not the case for the specifier (SU), as illustrated here on the basis of English and French.

- (14) (a) book-reading **English**
 (b) *student-reading of books
 (c) book-reader
 (d) *student-reader of books
- (15) (a) ouvre-porte **French**
 open-door
 ‘door-opener’
- (b) *ouvre-clef
 open-key
 ‘key-opener’

Construction-specific conditions have been proposed to account for the restrictions illustrated above, including the First Sister Principle (Roeper–Siegel 1978).

³ For an account of the difference in linear order of the parts of deverbal compounds see Di Sciullo (1996).

More general conditions have been evoked within GB theory to account for cases such as (14d) in terms of a Theta-Criterion violation. However, the latter solution does not extend to cases such as (14b), where the same argument is not saturated twice within the compound, viz., by the nominal suffix and by the nominal non-head.

The exclusion of the specifier (SU) from the derivation of compounds follows from the Asymmetry Hypothesis and the local domain of an asymmetrical relation, given the assumption that the subject is outside of the V domain, contrary to the direct object (DO), as in Chomsky (1998, 10).

(16) [SU [_v [V DO]]

If grammatical relations are asymmetrical relations in local domains, it follows that the derivation of deverbal compounds may only give rise to configurations where a complement, but not a specifier (SU), is the non-head of the configuration; the complement is the closest r-related position with respect to a head, the specifier (SU) does not qualify as such.

However, the specifier (SU) can be part of a deverbal compound as an agentive/instrumental suffixal head. The suffix is overt in languages such as English, Modern Greek, Hindi, Slavic and covert in Romance and in African languages. The relevant difference is depicted below.

(17) [[_v N V] -af]

(18) [[_v V N] - ϕ]

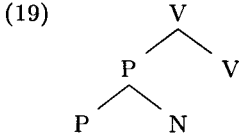
A theory based on asymmetrical relations applying in local domains is able to account for the restriction on the derivation of compounds without construction-specific conditions. The First-Sister Principle, as well as the Theta-Criterion are both too weak and too strong to account for the facts.

2.1.2. Romance verbs

There is evidence to the effect that the configurational properties of Romance denominal verbs, such as Portuguese *ensaboar* 'to rub soap on something/one' and *engarrifar* 'to put something in a bottle', as well as Italian *insaponare* and *imbottigliare*, are also organized in terms of asymmetrical relations.

We assume the following minimal structure for these expressions, as discussed in Di Sciullo (1997) on the basis of French and Italian. The verbal head

is a suffix, the prefix is prepositional, and the base form is nominal. The base noun is a complement of the prepositional prefix and the prepositional complex is an adjunct to the verbal head. The prepositional prefix and the base noun are in an $r_{predicate}$ relation and the prepositional complex is in $r_{modifier}$ relation with the verbal head.



This structure supports productive denominal verb formation in Romance languages, as most verbs in such languages are derived from a nominal (or an adjectival) base. This is not so for other languages, such as English for example, where the class of verbal suffixes is very restricted and where conversion is used for noun/verb alternations.⁴

⁴ We have shown elsewhere (Di Sciullo 1997) that the differences in verbal prefixation amongst Romance languages, exemplified here in (i), follow from the theory.

- (i) (a) **insaponare, impanare** **Italian**
 (b) **savonner, paner** **French**
 (c) **insaboar, impanar** **Portuguese**
 (d) **jabonar, panar** **Spanish**
 ‘to soap’, ‘to coat with bread crumbs’

The difference in prefixation between Italian and Portuguese on the one hand, and French and Spanish on the other, as well as differences between these two sets of languages with respect to the licensing of directional particles and directional resultatives follows from a difference in the projection or non-projection of the directional/prepositional feature of Aspect.

- (ii) (a) **buttare via il vaso/buttare il vaso via** **Italian**
 (b) **jeter le vase définitivement** **French**
 ‘throw the vase definitively’
 (c) **jogar fora o copo/jogar o copo fora** **Portuguese**
 (d) ***botar lejos el vaso/botar el vaso lejos** **Spanish**
 ‘throw away the vase/throw the vase away’
 botar el vase definitivamente
 ‘throw the vase definitively’
- (iii) (a) **Il fiume serpeggia al mare** **Italian**
 (b) **Le fleuve serpente *(jusqu’) à la mer** **French**
 (c) **O rio serpenteia o mar** **Portuguese**
 (d) **El río serpentea *(hasta) el mar** **Spanish**
 ‘The river snakes (its way) to the sea.’

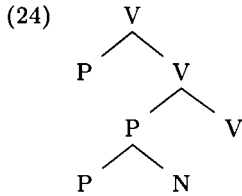
- (20) (a) **inscatolare, imbottigliare** **Italian**
 (b) **emboîter, embouteiller** **French**
 (c) **encaixar, engarrafar** **Portuguese**
 (d) **encajar, embotellar** **Spanish**
 'to box', 'to bottle'

Here again, the non-head cannot be the specifier SU of the verbal head, it can be the DO or the IO, as with *insaponare* and *inscatolare*, where the noun included in the derived verb is the DO in the first case and the IO in the second. Thus, as predicted by the Asymmetry hypothesis, DO may enter into a local asymmetrical relation with the head, but not SU. This holds for the languages under consideration notwithstanding the parametric variation with respect to the presence or the absence of a prefix. The examples in (21) illustrate this point.

- (21) (a) **im-person-ific-are** **Italian**
 (b) ***im-person-are**
 (c) **person-ifi-er** **French**
 (d) ***person-er**
 (e) **to person-ify** **English**
 (f) ***to person**

The Asymmetry Hypothesis also covers the combinatory restrictions amongst prefixes. We have shown in Di Sciullo–Klippel (1994) that in verbal structures, external prefixes, such as iterative and inverse prefixes, must precede internal directional and locational prefixes. The external vs. internal prefix distinction is exemplified in (22) and (23) and represented in (24), where external prefixes are outside the V DO domain, which includes internal prefixes.

- (22) (a) **réapporter, *areporter** **French**
 'to bring back again'
 (b) **réenlever, *enrelever**
 'to take off again'
- (23) (a) **to reencourage, *to encourage** **English**
 (b) **to reimprison, *to imprison**



As expected, clustering is possible amongst external prefixes, which are in $r_{adjunct}$ relation, but not amongst internal prefixes, which are in a r_{compl} relation.

- (25) (a) rereporter, *aemporter **French**
 'to report again', 'to bring to at'
- (b) reretacher, *aentacher
 'to restrain again', 'to stain to at'
- (c) to reunbutton, *to adetach **English**
- (d) to unrebutton, *to deattach

This result is independently supported by the observation that while there is no restriction on multiple adjunction, there is a unicity restriction on complementation (Grimshaw 1990; among others).

2.2. Summary

The Asymmetry Hypothesis makes the correct predictions with respect to the restrictions on morphological configurations, be they compounds or derived verbs. Moreover, it does so in a unified way, via the notion of asymmetry applying in local domains. In the next section, we show that our hypothesis also covers the properties of morphological paradigms.

3. Morphological paradigms

Morphological paradigms, such as category-changing affixes, are not unorganized lists of morphemes, as the following properties hold for their members.

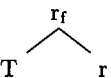
- (26) (a) They are limited in number.
 (b) Their composition with roots is restricted.

These properties are not those of lexical elements, such as nouns and verbs, as the elements of lexical classes are not limited in number and their composition with other categories is dependent on the properties of functional categories. We claim that the properties of morphological paradigms follow from the very architecture of these paradigms, which is based on asymmetrical relations.⁵

In effect, one consequence of our hypothesis is that morphological paradigms m are not unorganized set of elements, rather they are organized in terms of pairs of abstract grammatical elements x, y, \dots in asymmetrical relations. Actual morphemes are part of a given paradigm only if they are licensed by such relations. Let us assume the following:

- (27) For each x in m there must be a y such that for each (x, y) there is an r such that $r(x, y)$ is true in m .

Typically, morphological paradigms are constituted of formatives which are the morphological spell-out of grammatical features. Thus, for example, Case paradigms in languages such as Rumanian and Modern Greek, specify case features for nouns. Likewise, the inflectional paradigms specify mood, tense as well as person and number features for verbs. Here again, the specification goes from types of categories to restrictors, as the affixes generally differ as to whether they specify inflectional features of verbs or nouns. Let us take an asymmetrical relation (r_f) to hold between types of categories (T) and restrictors (r) such that there is a unidirectional relation that goes from types of categories to restrictors. This local asymmetrical relation can either be expressed in terms of binary branching structure, as in (28a), or in terms of pairs of elements, as in (28b):

- (28) (a)  (b) $r_f(T, r)$

We thus take morphological paradigms to be defined in terms of asymmetrical relations over pairs of grammatical elements, viz., categories and restrictive features. No formative, actual affix, can be part of a paradigm if not licensed by an asymmetrical relation. We consider the properties of inflectional and derivational paradigms in what follows.

⁵ Inflectional paradigm economy has been discussed in Carstairs (1983) in terms of a tendency of keeping the total of paradigms for any word-class close to the logical minimum. Our hypothesis presents a rationale for why this must be the case.

3.1. Inflectional paradigms

Let us start by considering a simple case, the nominal phi-feature paradigm of Italian. The forms of the paradigm are listed in (29) and the examples in (30) illustrate their distribution. Thus, in Italian each pair of inflectional features for nominal categories has a morphological spell-out.

- (29) N fem masc
 sing: -a -o
 plur: -e -i

- (30) (a) rosa, rose
 'rose_{fem,sing}', 'rose_{plur}'
 (b) libro, libri
 'book_{masc,sing}', 'book_{plur}'

According to our proposal, inflectional affixes are defined in terms of categories and restricting features, and each formative of inflectional paradigms is licensed via an asymmetrical relation that goes from a type of category, here N, to a restricting inflectional feature, here the number feature. Let us posit the following.

- (31) r_{phi}
 Categorical types include phi restrictors.

The local asymmetrical approach to the architecture of morphological paradigms allows us to articulate the inflectional features for nominal categories, as in (32).

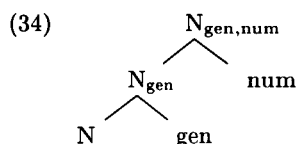
- (32)
$$\begin{array}{c} N_{\text{phi}} \\ \swarrow \quad \searrow \\ N \quad \quad \text{phi} \end{array}$$

In our framework, r_{phi} is an asymmetrical relation as it unidirectionally relates nominal types of categories N and A to phi-features, restrictors such as gender features (gen) and number features (num). Thus, we have the following r_{phi} for Italian:

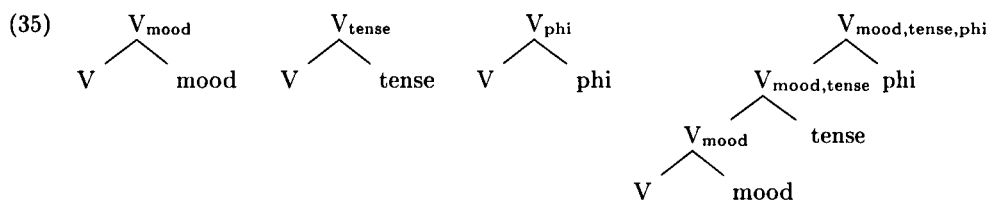
- (33) $r_{\text{phi}}(N_{\text{masc}, \text{sing}})$:-o
 $r_{\text{phi}}(N_{\text{fem}, \text{sing}})$:-a
 $r_{\text{phi}}(N_{\text{masc}, \text{plur}})$:-i
 $r_{\text{phi}}(N_{\text{fem}, \text{plur}})$:-e

Our theory makes the correct prediction with respect to the size of the nominal inflectional paradigm. The inflectional affixes are restricted to four morphemes as each affix is the morphological spell-out of r_{phi} applied to nominal categories.

The representation in (34) depicts the compositionality of the nominal inflectional features projection. These relations are not formally expressed by standard representations such as in (29) above.



Our proposal also covers inflected categories such as verbs, the phi-features of which include person, tense and mood features, which we will not discuss here. The configuration below represents the dependencies among inflectional features. Languages may vary with respect to the morphological spell-out of the inflectional features, as well as with respect to the linear order of the actual affixes with respect to the verb.



In our view, there can be no affix that is part of an inflectional paradigm that is not licensed by r_{phi} , even though the same morpheme may instantiate more than one asymmetrical relation.⁶

⁶ We thank an anonymous reviewer for pointing out to us that, in Hungarian, the same phi-morpheme can be part of both nominal and verbal paradigms for many person/number combinations.

(i) fut-UNK vs. ut-UNK (ii) kap-JA vs. csap-JA
 run-1pl way-1pl get-3sg.def-obj tap-3sg
 'we run' 'our way' 'he gets' 'his tap'

3.2. Derivational paradigms

Derivational paradigms are defined in terms of formal features, such as categorial features, (Williams 1981; Di Sciullo–Williams 1987), as well as semantic features, such as argument structure features (Di Sciullo 1995) and aspectual features (Di Sciullo 1997). The following sections present evidence to the effect that the architecture of derivational paradigms is based on local asymmetrical relations ranging over formal and semantic features.

3.2.1. Formal features

Derivational affixes, such as *-er* and *-able*, have categorial features, as they participate in category-changing operations. In our theory, the paradigm of category changing affixes is defined in terms of asymmetrical relation between pairs of grammatical elements. Let us posit the following lexical relation r_{lex} .

- (36) r_{lex}
 Categorial types include lexical restrictors.

We will take the lexical categories to be restricted to the lexical features N and V, as in Chomsky (1970; 1995). Assuming a privative feature notation, we obtain the following definitions for the lexical categories.

- (37) categories: $\frac{V}{V} \mid \frac{N}{N} \mid \frac{A}{V} \mid \frac{P}{N}$

We will thus assume that the asymmetrical relation r_{lex} ranges over the formal features N and V of both the category types and the restrictors, as in the following representation, where types of categories are in capitals, but not restrictors.

- (38)
-
- ```

 graph TD
 r_lex --- T
 r_lex --- r
 T --- N
 T --- V
 r --- n
 r --- v

```

The following catalogue includes morphological spell-outs that result from the asymmetrical relations  $r_{lex}$ , which go from the categorial type (of affixal heads) to the feature restrictors (of the non-heads).

- (39) (a)  $T_N r_n$ -----  
 $r_{lex}$  (N,n/v): -té beauté  
 'beauty'  
 (N/V,n): -el formel  
 'formal'
- (b)  $T_N r_v$ -----  
 $r_{lex}$  (N,v): -ion destruction  
 'destruction'  
 (N,n/v): -ance aisance  
 'wealthiness'  
 (N/V,v): -able adorable  
 'adorable'
- (c)  $T_V r_n$ -----  
 $r_{lex}$  (V,n): -er boutonner  
 'to button'  
 (V,n/v): -ir pâlir  
 'to become pale'
- (d)  $T_V r_v$ -----  
 $r_{lex}$  (V,n/v): -izer formalizer  
 'privatize'  
 (V,n/v): -ifier bonifier  
 'bonify'  
 (N/V,v): -eur enchanteur  
 'enchanting'

Thus, (39a) includes morphemes that result from the application of  $r_{lex}$  from N types of affixes to n restrictors; (39b) includes morphemes that result from the application of  $r_{lex}$  from N types of affixes to v restrictors, and so on. Different morphemes may instantiate the same  $r_{lex}$  relation, even though they will differ with respect to semantic asymmetrical relations, as exemplified in (39a) and (39d), as well as in (40). Moreover, the same morpheme may instantiate more than one  $r_{lex}$  relation, as evidenced in (41).

(40) (a) aim-able  
 'loveable'

(b) pétill-ant  
 'sparkling'

(41) (a) jouv-et  
 play-small  
 'toy'

- (b) livre<sub>N</sub>-et  
 book-small  
 'booklet'

The asymmetrical type/restrictor relation captures the articulation of the paradigm, as well as the c-selection of the affixes. Furthermore, this system predicts that there is no productive category shifting with prepositions, as the latter have no nominal or verbal features.

### 3.2.2. Semantic features

Derivational affixes are also associated with semantic features. We focus on their aspectual features here. We have shown elsewhere on the basis of French and Italian (Di Sciullo 1997), that prefixes such as *a-*, *en-/in-* and *re-/ri-* may affect the aspectual structure of the verbal projection they are a part of. The aspectual contribution of the prefixes differ whether they are external or internal to the minimal argument structure projection of the verb. Internal prefixes, such as *a-* and *en-* may affect the terminative reading of the event, providing an endpoint to an unbounded event. They may change Activities onto Accomplishments, as exemplified with the French and Spanish examples in (42). This is not so for the iterative prefix, which affects a bounded event, viz., Achievements, and Accomplishments, without affecting the internal structure of the event, as exemplified in (43) and (44) respectively. The difference in aspectuality may also be spelled out by different verbs, as it is the case in Spanish, where similar facts are observed.

- (42) (a) Il a couru pendant une heure/\*en une heure. French  
 Corrió durante una hora/\*en una hora. Spanish  
 'He ran for an hour/in an hour.'
- (b) Il est accouru \*pendant une heure/en une heure.  
 Acudió \*durante una hora/en una hora.  
 'He rushed up for an hour/in an hour.'
- (43) (a) La bombe a (ré)explosé \*pendant cinq secondes/en cinq secondes. French  
 La bomba (re)explotó \*durante cinco segundos/en cinco segundos. Spanish  
 'The bomb (re)exploded for five seconds/in five seconds.'
- (b) \*La bombe a arrêté de (ré)exploder.  
 \*La bomba paró de (re)explotar.  
 'The bomb stopped (re)exploding.'

- (44) (a) Il a (re)construit la maison \*pendant un an/en un an. French  
 (Re)construyó la casa \*durante un año/en un año. Spanish  
 'He (re)built the house for a year/in a year.'
- (b) Il a arrêté de (re)construire une maison.  
 Paró de (re)construir una casa.  
 'He stopped (re)building a house.'

These facts are expected given the spatio-temporal theory of aspect (Hale 1984; Kipka 1990; Tenny 1994, and related works), according to which spatial categories participate in aspectual structure and contribute to determine if an event is terminative, i.e., if it has a natural endpoint, or not.

We assume that aspect is defined in terms of the aspectual features terminative [T] and sub-interval [S], along the lines of Di Sciullo (1997), where natural classes of aspectual categories share combinatorial properties. Verbs have aspectual features: States: [-T,-S] (*to know, to love*), Achievements: [+T,-S] (*to explode, to find*), Activities: [-T,+S] (*to bring, to run*); Accomplishments: [+T,+S] (*to construct, to destroy*). Prepositions also have aspectual features. They contribute to aspectual structure as they are relational categories establishing the spatio-temporal properties of an event or a situation. Reformulating the proposal in terms of absolute values for the aspectual features, we obtain:

- (45) categories: activity | accomplishment | achievement | state  
 features:       s           s                   t           t

We propose here a finer-grained articulation of aspect features, that distinguishes the aspectual contribution of the following types of semantic categories, space (S) and time (T) categories, with respect to the restrictive aspectual features sub-interval (s) and terminus (t).

- (46)
- $$\begin{array}{c}
 r_{asp} \\
 \swarrow \quad \searrow \\
 T \qquad \qquad r \\
 \swarrow \quad \searrow \quad \swarrow \quad \searrow \\
 S \qquad T \quad s \qquad t
 \end{array}$$

We posit the following asymmetrical relation:

- (47)  $r_{asp}$   
 Categorial types include aspectual restrictors.

$r_{asp}$  organizes the aspect-changing affix paradigm. Each formative of the paradigm is the morphological spell-out of an asymmetrical relation holding between spatio-temporal types of categories and aspectual feature restrictors. The catalogue for aspect-changing prepositional prefixes (P) includes the following.

- (48) (a)  $r_{asp}$  ( $P_T$ , s/t): re- recoller  
'to reglue'  
dé- décoller  
'to unglue'
- $r_{asp}$  ( $P_T$ , t): pré- prédater  
'to predate'  
post- postdater  
'to postdate'
- $r_{asp}$  ( $P_T$ , s): ex- exposer  
'to expode'  
im- imposer  
'to impode'
- (b)  $r_{asp}$  ( $P_S$ , s/t): é- émettre  
'to emit'  
trans- transmettre  
'to transmit'
- $r_{asp}$  ( $P_S$ , t): sous- sousestimer  
'to underestimate'  
sur- surestimer  
'to overestimate'
- $r_{asp}$  ( $P_S$ , s): a- apporter  
'to bring to'  
en- emporter  
'to bring away'

Thus, *re-* and *dé-* are temporal prepositional morphemes ( $P_T$ ) and they are restricted by the aspectual feature (s/t); *pré-* and *post-* are also of the same type  $P_T$  and they are restricted by the terminative feature (t), they range over events with a terminus and no sub-internal; *ex-* and *in-* are also of the type  $P_T$  and they range over events with a sub-interval (s). The prefixes *é-* and *trans-* are of the spatial type  $P_S$  and they are restricted by the aspectual feature (s/t); *a-* and *en-* are of the same type  $P_S$  and they are restricted by the sub-interval feature (s); *sous-* and *sur-* are also of the type  $P_S$  and they are restricted by the boundedness feature (t). The aspectual differences are evidenced by examples such as the ones in (49) and (50) where a sub-interval interpretation of the event is forced.

- (49) (a) Il n'a pas fini de recoller les morceaux.  
'He did not finish to reglue the pieces.'
- (b) \*Il n'a pas fini de postposer la rencontre.  
'He did not finish to postpone the meeting.'
- (50) (a) Il n'a pas fini de transmettre le message.  
'He did not finish to transmit the message.'
- (b) \*Il n'a pas fini d'apporter le document.  
'He did not finish to bring the document.'

Here again, even though singular morphemes may instantiate more than one asymmetrical relation, as it is the case for *re-*, which is iterative in *to rewrite* and intensive in *to reassure*, no singular morpheme can be licensed without being the morphological spell-out of an asymmetrical relation.

Assuming the Adjunct Identification Condition (Di Sciullo 1997), according to which an adjunct identifies an underspecified feature of the category it adjoins to, our hypothesis accounts for fine-grained combinatorial properties of prepositional prefixes with verbs. We correctly predict that aspect changing spatial ( $P_s$ ) prefixes, such as French *a-* and *en-* may not combine with bounded events, [t] verbs (Accomplishments and Achievements), such as French *construire* 'to construct' and *peindre* 'to paint', *trouver* 'to find' and *exploser* 'to explode', a fact not predicted by c-selection, as c-selection does not range over aspectual features. Thus, fine-grained semantic restrictions hold for prefixes in their combinations with verbs. They follow from the asymmetrical relation organizing aspect changing affixal paradigm in conjunction with independently motivated conditions of the grammar.

### 3.3. Summary

Asymmetrical relations account for the properties of derivational paradigms, which are complex objects ranging over formal and semantic features. We proposed that the structure of these paradigms is organized and restricted by such relations. Our proposal accounts for the properties of affixal paradigms, their restrictiveness in number and their combination with roots, in the following unified way.

- (51) (a) Derivational paradigms are limited in number as their members are the morphological spell-out of asymmetrical relations ranging over a limited set of types of categories and restrictive features.

- (b) Fine-grained combinatorial restrictions hold for the members of morphological paradigms, as they are the spell-out of categorial and semantic asymmetrical relations.

#### 4. Theoretical consequences

A first consequence of our proposal is for the theory of grammatical relations. It posits the existence of a basic asymmetrical relation that provides a purely relational definition of morphological configurations and morphological paradigms, and thus derives the effects of rule based theories of morphology.

A second consequence of our proposal is for the theory of paradigms. We provided evidence to the effect that the geometry of morphological paradigms is based on asymmetrical relations, restricting the combinations of affixes with roots, and deriving the effects of c-selection and s-selection.

A third consequence is for the theory of economy of the grammar. Our proposal brings support to a unified organization of the grammar, where a central property, asymmetry, plays a role in different sub-systems. It covers the configurational properties of morphological objects, compound and derived forms in a unified way, and it extends to the organization of morphological paradigms.

#### References

- Carstairs, A. 1983. Paradigm economy. In: *Journal of Linguistics* 19: 115–25.
- Chomsky, N. 1970. Remarks on nominalization. In: Jakobs, R. – Rosenbaum, P. (eds): *Readings in English transformational grammar*. Mouton, The Hague.
- Chomsky, N. 1981. *Lectures on government and binding*. Foris, Dordrecht.
- Chomsky, N. 1995. *The Minimalist Program*. The MIT Press, Cambridge MA.
- Chomsky, N. 1998. *Minimalist inquiries*. Ms. MIT.
- Collins, C. 1997. *Local economy*. The MIT Press, Cambridge MA.
- Di Sciullo, A.M. 1995. X' selection. In: Roorick, J. – Zaring, L. (eds): *Phrase structure and the lexicon*, 77–107. Kluwer, Dordrecht.
- Di Sciullo, A.M. 1996. Modularity and X<sup>0</sup>/XP asymmetry. In: *Linguistic Analysis* 26: 3–28.
- Di Sciullo, A.M. 1997. Prefixed-verbs and adjunct identification. In: Di Sciullo, A.M. (ed.): *Projections and interface conditions*, 52–74. Oxford University Press, Oxford.
- Di Sciullo, A.M. – Klipple, E. 1994. *Modifying affixes*. WECOL XXIII. University of Washington, Seattle.



- Di Sciullo, A.M. – Williams, E. 1987. *On the definition of the word*. The MIT Press, Cambridge MA.
- Grimshaw, J. 1990. *Argument structure*. The MIT Press, Cambridge MA.
- Hale, K. 1984. *Notes on world view and semantic categories: some Walpiri examples*. Ms. MIT, Cambridge MA.
- Hale, K. – Keyser, J. 1993. *The view from Building 20. Essays in linguistics in honor of Sylvain Bromberger*. The MIT Press, Cambridge MA.
- Hornstein, N. 1995. *Logical Form*. Blackwell, Oxford & Cambridge MA.
- Kayne, R. 1984. *Connectedness and binary branching*. Foris, Dordrecht.
- Kayne, R. 1994. *The antisymmetry of syntax*. The MIT Press, Cambridge MA.
- Keyser, J. – Roeper, T. 1995. *Anti-symmetry and leftward movement in morphology*. Ms. UMASS and MIT.
- Kipka, P. 1990. *Slavic aspect and its implications*. Ph.D. dissertation, MIT, Cambridge MA.
- Reinhart, T. 1983. *Anaphora and semantic interpretation*. Croom Helm, London.
- Reinhart, T. 1995. *Interface strategies*. OTS Working Papers. Utrecht University, Utrecht.
- Roeper, T. – Siegel, M. 1978. *A lexical transformation for verbal compounds*. In: *Linguistic Inquiry* 9: 199–260.
- Tenny, C. L. 1994. *Aspectual roles and the syntax-semantics interface*. Kluwer, Dordrecht.
- Williams, E. 1981. *On the notions lexically related and head of a word*. In: *Linguistic Inquiry* 12: 245–74.
- Williams, E. 1994. *Thematic structure in syntax*. The MIT Press, Cambridge MA.

Address of the author: Anna-Maria Di Sciullo  
Université du Québec à Montréal  
Case Postale 8888, succursale Centre-Ville  
Montréal (Québec) H3C 3P8  
Canada  
di\_sciullo.anna-maria@uqam.ca



PRODUCTIVITY IN WORD FORMATION (WF):  
A MORPHOLOGICAL APPROACH\*

WOLFGANG U. DRESSLER AND MÁRIA LADÁNYI

**Abstract**

Grammatical productivity of the potential system of WF (1) is distinguished from type frequency (on the level of language as institution) and token frequency (level of performance), where pragmatic factors intervene (2). It is a hyponym of legal regularity and creativity, different from default, probability of WFR application, e.g., in rule competition, and analogy. Restrictions on actual vs. potential domains do not restrict productivity itself, as shown by ordinal number formation (3) and German feminine motion (4). The main part focusses on qualitative gradation of grammatical productivity (5), its theoretical bases and methodological consequences (6), including stratified lexica (7). External evidence from poetic license is adduced in 10. Hierarchical paradigmatic organization of WF (8) includes the levels of derivational paradigm, microclass, class, and their higher sets. Typologically, agglutinating WF (e.g., Hungarian) is richer than inflecting WF (9). This contribution represents a plea (a) for the analytic separation of grammar and pragmatics also in the area of productivity, (b) for the unity of inflection and derivation, (c) for the close connection between morphological meaning and form.

**1. Introduction**

Morphological productivity, on the level of the potential system, can still be defined in Schultink's way (translated by van Marle 1985, 45) as "the possibility for language users to coin, unintentionally, a number of formations which

\* A first version of this contribution has been republished as Dressler – Ladányi (1998). Both versions are outcomes of our common project on word formation theory and typology (contrasting WF in inflecting and agglutinating languages), supported by project 32öu15 of the Stiftung Aktion Österreich-Ungarn (Osztrák-Magyar Akció Alapítvány) and by the Telegdi Zsigmond Foundation (Ladányi for 1997/98).

are in principle uncountable.”<sup>1</sup> We want to follow this direction, which we think has been given up explicitly or implicitly by too many of our colleagues. Thus the subtitle of this paper “a morphological approach” is provocative, as it implies that some other current approaches to morphological productivity are not truly morphological, i.e., in the sense of morphology as morphological grammar. For example, connectionist and related analogical approaches can be said to be basically lexical, approaches within word syntax are basically syntactic, in split morphology the basic unity of morphological productivity in both WF and inflection is excluded. Then there are approaches to morphological productivity which are essentially computational or psycholinguistic. Thus the great majority of theoretical and descriptive work on WF productivity is either not in terms of a theory of morphological grammar or of a theory which either splits derivational morphology from inflection or splits morphological meaning (morphosemantics) from morphological form (morphotactics), as in Beard’s (1995) separation thesis. We will argue against these views and try to show what we mean by a truly grammatical-morphological approach to morphological productivity.

We assign our concept of WF productivity to the level of grammatical WF competence which accounts for potential words (in the sense of Aronoff 1976). This level is close to Coseriu’s (1975) notion of language as potential system and to Chomsky’s (1986) notion of internal language. This level has to be sharply distinguished, at least analytically, from external language, or more precisely, language as an institutional system of norms. Here belong both actual and accepted words, and grammatical productivity is realized on this second level as actual type frequency of complex words derived via the same WFR. On the third level of performance, the concept which corresponds to productivity on the first level, is token frequency. Therefore we are against deriving productivity from type and/or token frequency (Baayen 1992; Baayen-Renouf 1996; cf. van Marle’s 1992 criticism).

Before we continue with the architecture of this tripartite hierarchy, we want to stress that we are going to deal only with grammatical WF productivity as opposed to extragrammatical morphological productivity. Extragrammatical operations of so-called expressive morphology, for example abbreviations or echo-word formation, may be formed via analogical lexical relations, such as E *zigzag*, H *csitt-csatt* = G *klippklapp*, F *froufrou* (cf. Mayerthaler 1977;

<sup>1</sup> Also cited by Baayen – Lieber (1991, 808). Their research program, of arriving finally at measures of productivity at the levels of norm and of performance, is basically also ours. Here, however, we concentrate on the level of the potential system.

Dressler–Merlini Barbaresi 1994, 40). They are not subject to grammatical competence and behave differently from WFRs in many ways and thus remain outside our discussion, although the distinction is of considerable theoretical interest, insofar as, for example, Anderson’s (1992) “a-morphous” morphology appears to differentiate between grammatical and extragrammatical morphology much less than conceptions in terms of directional WFRs. We do not insist on the rule-format (e.g., of the Aronoff type) for describing productive grammatical patterns, but use it as a very transparent device for characterizing them.

## 2. Conceptions and misconceptions of productivity

**2.1.** Morphological productivity should not be confused with, or reduced to (Poitou 1997), its hyperonym regularity (=homogeneity of input-output relations of rules): all productive patterns are rule-governed and thus their morphosemantic and morphotactic motivation is systematically analysable in a rule format (cf. Kastovsky 1982, 157). But, by definition, also unproductive rules are regular, both in inflection and derivation. Thus, for example, Berrendonner and Clavier (1997) describe the regularity of two French types of suffixation of *-age*, one being productive, the other one unproductive but still regular. Or German denominal adjective formation via the suffix *-ig* is productive, whereas deverbal *-ig* adjective formation is not. Nevertheless we assume that productivity, understood as rule-governed creativity (cf. Bauer 1983, 63), is a prototypical property of rules. In other words, productivity is a hyponym of both creativity and regularity, whereas rule-changing creativity is not regular—note the great variety and idiosyncrasy of the most creative poetic occasionalisms (cf. also Rainer 1993, 31f).

**2.2.** Next, such as in inflectional morphology, productivity should not be confused with the default status (as in Clahsen et al. 1996). For example, in Hungarian denominal verb formation, suffixation with both *-(o/e/ö)z-* and *-(o/e/ö)l* is productive, but *-(o/e/ö)z-* is the default, because its domain of application is less restricted: the suffix *-(o/e/ö)l* can be attached productively only to consonant-final roots (except *-l/r-*), which are mostly monosyllabic, plus English verbs and action nouns in *-ing*,<sup>2</sup> as in the loan words:

<sup>2</sup> This is a lexical restriction. For, if the base is an English loan in *-ing* which is not an action noun, such as *puding*, then the denominal verb is *puding-oz(ik)* not *\*puding-ol*.

- (1) szörf-öl = szörf-öz-ik 'surf', lízing-el 'lease', sopping-ol 'shop', szlalom-oz-ik 'slalom', videó-z-ik 'watch/make a video recording'

Such restrictions do not hold for denominal suffixation with  $-(o/e/ö)z-$ .

As to the semantic aspect of WFRs, our model, Natural Morphology,<sup>3</sup> has always consistently distinguished word formation meaning and word meaning (cf. Rainer 1993, 132–5), i.e., the morphosemantics of grammar and the lexicon.<sup>4</sup> Only WF meaning is directly relevant to WF productivity.

**2.3.** Grammatical WF productivity accounts for what is a potential or possible legal complex/derived word, but only partially for what is a probable word, i.e., for how probable it is that a potential word is produced (on the level of performance) or accepted (on the level of language as institution/norm). Thus we cannot fully agree with Kirscher-Durand's (1997, 178) claim "le lexique potentiel garantit la vraisemblance—et non l'occurrence—des signes linguistiques." Clearly, the occurrence of a word-formation is due to variables on the levels of norm and/or performance, and one might calculate from the set of relevant factors a certain probability of occurrence, but on the level of the potential system, only the possibility (legality) or impossibility (illegality) of a conceivable derived word can be accounted for, not its probability (despite the arguments in Rainer 1993, 32–4).

**2.4.** Grammar is especially involved in the various constellations of rule rivalry among competing rules, which may be either productive or unproductive. In partial recapitulation of the gradation of inflectional rule competition in Dressler (1997, 14–6) we can distinguish the following five constellations (a–e):<sup>5</sup>

<sup>3</sup> Cf. Dressler et al. (1987); Wurzel (1984); Kilani-Schoch (1988).

<sup>4</sup> For discovery procedures of how to find WF meanings via lexicological analyses of word meanings, cf. Corbin (1991); Temple (1996).

<sup>5</sup> The psycholinguistic basis, on the level of performance, is the hypothesis of Baayen (1989, 227): "in lexical access a rule-driven procedure operates in parallel with an item-driven access procedure." We may assume that the automatic application of a productive rule mechanism is very efficient, both in production and perception, in general more so than lexical retrieval of whole words (and especially of whole inflectional forms), provided that they are productively formed and thus entirely predictable, cf. also Baayen (1989, 4, 210ff, 220) and Frauenfelder – Schreuder's (1992) race model. Baayen and Neijts (1997), however, think that only low-frequency words and neologisms are produced and understood by rule.

- (a) One polar case is the constellation of one fully productive WFR which does not compete with any other WFR. Since accepted outputs of a (even fully productive) WFR are stored in the lexicon, their existence may influence the application of even a fully productive WFR (e.g., in case of blocking). Thus also surface analogies cannot be excluded, as in the case of G *ent-drei-t* 'divided into three', formed with the productive WFR of forming verbs with the prefix *ent-* 'dis-'. But in the act of performance of the textual sequence *ent-zwei-t*, *ent-drei-t* 'divided into two, divided into three' (poem by J. Ringelnatz), the ludic occasionalism *ent-drei-en* represents a surface analogy to the immediately preceding past participle of the existing word *ent-zwei-en*.
- (b) The opposite pole is represented by the constellation of the absence of any WFR, e.g., in the case of a word which belongs to an isolated paradigm, particularly if the respective form belongs to an idiosyncratic part of the paradigm. Here only lexical retrieval of the full derivational form is possible. An example is G *ein-sam* 'alone' (2.6).
- (c) A constellation which is near this pole (b), is represented by the constellation of an unproductive WFR. Such rules have still the function of motivation, although not of lexical enrichment (cf. Baayen 1989, 225f). Since poetic occasionalisms are more frequent with unproductive WFRs than with non-rules (constellation (b)), we cannot fully agree with Baayen (1989, 193): "Given that unproductive rules have no psycholinguistic reality, unproductive formations wholly depend on accurate memory retrieval" (cf. Pinker-Prince 1994, 327). Clearly, lexical retrieval is the default operation, but reference to a schema or unproductive WFR is not excluded, as a secondary motivation, in conscious neoformations. For example, note the conscious nationalistic revival of fossilic Croatian WFRs under the puristic, Anti-Serbian language policy of the government of newly independent Croatia in the nineties.
- (d) Another relevant constellation is represented by the presence of a slightly productive WFR: the efficiency of the rule mechanism in its potential domain is presumably very limited, constant lexical checking may be necessary whether the complex item perceived or to be produced really exists or is adequate in the given circumstances. Thus the rule may be only rarely used in processing, particularly when more "conscious" efforts are called for, as when processing puns, new words, nonsense words, or in cases of misunderstanding, in learning situations, in evaluations of forms. Such rules may then serve as fall-back procedures (cf. Baayen 1989, 212, 221ff; Frauenfelder-Schreuder 1992, 170; Sandra 1994, 245f).

- (e) Another important constellation consists in competition (rivalry) of equally productive WFRs: here lexical retrieval is necessary in production in order to decide which rule to apply on the level of language as norm and in performance. Thus this necessity is only relative, it might be cancelled if the speaker feels unbound by norms, as in the case of “abnormal” mental states (e.g., when alcoholized). In case that one of the competing rules is more productive or the default or at least the more general case, lexical look-up is more likely for the competing rule which represents the special case (cf. Baayen 1989, 14f). Only when the domains of the rival rules are complementary (disjunct, cf. Baayen 1989, 13f), is lexical look-up not necessary.

Obviously, the assignment of WFRs to these five constellations presupposes the establishment of degrees of productivity. This gradience must be established deductively on purely linguistic grounds (see section 5), because psycholinguistic considerations either belong to the level of performance or to the justification of universal preferences, and of course to the methodology of empirical, inductive testing.

2.5. However, the most important factors of probability are of a pragmatic nature which are definitely excluded from grammaticality, such as pragmatic need/usefulness or incompatibility/inconsistency, stylistic restrictions, such as:

- (a) lexical blocking (homophony and synonymy constraints, cf. Rainer 1988; Bauer 1983, 87f);
- (b) pragmatic incompatibilities of a referential nature (cf. Corbin 1997a, 87 and Bauer’s (1983, 85–7) nameability requirement);
- (c) lack of pragmatic need or usefulness in case of accidental lexical gaps (cf. van Marle 1992, 153f; Corbin 1997a). Here, the acceleration of modernization has led to an explosion of neologisms in many languages (cf. Petralli 1996, 18–22);
- (d) pragmatic self-evidence, therefore lack of need to name, cf. *beard-ed man* (not self-evident) vs. *\*ey-ed man* (cf. Kastovsky 1982, 160);
- (e) stylistic restrictions (cf. Rainer 1993, 125);
- (f) sociolinguistic factors (cf. Alloni-Feinberg 1977).

However, existing idiosyncratic words, which are formed consciously but illegally (cf. Christofidou 1994; Dal 1997) are outside the potential domain of a WFR. The example we give comes from a highly successful slogan of the Austrian tourist office (portraying hikers in an idealized mountainous region):



## (2) Wander-bares Österreich 'hike-able Austria'

where \**wander-bar* comes, syntagmatically, in an illegal way, from the intransitive (one-place) base verb *wander-n* 'to hike', but paradigmatically, via surface analogy, from *wunder-bar* 'wonder-ful'. Partial acceptance of this new word is due to its (consciously planned) connotative evocation of its model, which is guaranteed by the double relation of a phonological minimal pair and of a superficially parallel WF.

A Hungarian example which may illustrate three of the above factors (a, c, d) follows: Kiefer (1998) has claimed that productive noun incorporation into Hungarian complex predicates, i.e., juxtapositional noun-verb constructions, is only possible if the complex activity referred to is institutionalized, which explains why (3a) is well-formed, whereas (3b) is not:

- |                                                                   |                                                                     |
|-------------------------------------------------------------------|---------------------------------------------------------------------|
| (3) (a) beteg-et látogat<br>sick-acc visit<br>'(s)he sick-visits' | (b) *barát-ot látogat<br>friend-acc visit<br>'(s)he *friend-visits' |
|-------------------------------------------------------------------|---------------------------------------------------------------------|

The reason is that visiting patients in hospitals or at their homes is an institutionalized, ritualized social activity, whereas visiting friends is not. But if we think of the 19th century aristocratic ritual of visiting respected acquaintances on a fixed day each week for tea, then by vividly depicting the pragmatics of such rituals, the use of (3b) becomes acceptable, i.e., an institutionalized, ritualized social activity is recreated, and Bauer's (1983, 85–7) nameability requirement is fulfilled. The fact that this activity has not been named with (3b), is due to lexical blocking by the, now nearly obsolete, loan verb:

- (3) (c) vizit-ál (integrated with the H suffix
- ál*
- )

Thus Kiefer's insight should be reformulated in the following way: on the level of grammatical productivity, these complex predicates may not refer to individual acts with individual reference, but to general activities, and as to actual types on the level of social norms, reference must be made to a routine activity. Therefore (3b) is potential, thus legal or grammatical, but not acceptable under normal, current social conditions.

**2.6.** Finally, there is the important distinction to be made between (a) surface analogy vs. (b) analogy via a schema vs. (c) rule productivity:

- (a) in surface analogy (cf. Motsch 1981) an actual model is necessary, usually it is a very similar word (e.g., a rhyme word, cf. Thymé et al. 1994, 453ff).

When asked, the creator of an instance of surface analogy can (and often even spontaneously does) name the exact word which served as the actual model (cf. below);

- (c) with rule productivity, there is no necessity of an actual model, the rule-governed creation of a new word depends on the precise abstract pattern (describable in a rule format). When asked, the creator of the new word can name several words (of different types) formed by this rule;
- (b) in analogical creation via a schema (cf. Köpcke 1993), performance depends on prototypes identifiable as actual words, but no exact pattern can be described because of non-prototypical members of the schema. The creator's behaviour lies inbetween those of (a) and (c).

Clearly, rule productivity must be differentiated from surface analogy. A good example given by Motsch (1981) is: the isolated and idiosyncratic derivation of G *ein-sam* 'alone' from the numeral *ein-* '1' has inspired the poet Christian Morgenstern to form the occasionalism *zwei-sam* 'two alone' (referring to two lovers). Examples of accepted neologisms formed via surface analogy are F *al-lunir* 'to land on the moon' (from *la lune* 'moon') and (much earlier) *amerrir* 'to alight on water' (from *la mer* 'sea'): both neologisms have been formed from the one actual model of *atterrir* 'to land' (from *la terre* 'earth, land'); in fact the double orthographic ⟨rr⟩ of the earlier neologism has no other explanation than the influence of *atterrir*. Thus, surface analogies are paradigmatic creations. But often they are also due to syntagmatic crossreference to a word in the same text (for textual conditioning of neologisms cf. Matussek 1994).

**2.7.** Related to both **2.6** (a) and **2.6** (b) is loan transposition, due to analogy (often schematic) to another language. Rainer (1997b) has amply discussed the following example of conscious, schematic loan transposition: in analogy to pairs such as It *man-u-ale* 'manual' and its Latin source *manu-alis*, learned Italian adjectives may end in *-uale* rather than in *-ale*. Loan transposition may result in the rise of a (slightly) productive WFR, as is the case with the import of the Latin (incl. Neolatin) productive suffix *-anus* into Italian deanthroponymic adjective formation in *-ano*, e.g., in *Copernic-ano*, *Luter-ano* (Rainer 1996), cf. also section 7.

### 3. Potential and actual domains

**3.1.** In many models, productivity is not an all-or-nothing concept, but a gradual one. Usually the notion "degrees of productivity" refers to pragmatic

factors on the level of language as norm (cf. 2.5). For degrees of grammatical productivity much more relevant is Booij's (1977, 5) definition:

- (4) "the degree of productivity of a WF-rule can be seen as inversely proportional to the amount of competence restrictions on that WF-rule"

In our view, however, competence restrictions of this type do not limit productivity itself but only the potential domain of rule application.<sup>6</sup>

**3.2.** A case in point is ordinal numeral formation, which shows that productivity is best expressed as the ratio of potential words to conceivable words, independent of type and token frequency (cf. 3.7). Ordinal numerals, formed from cardinal numerals, represent a closed class, i.e., a closed actual domain of WFRs. Nearly all cardinal numerals are formed via composition of basic units. Typically, the most basic units are few, e.g., the numerals 1–9 plus irregular/suppletive compounds, such as E *eleven, twelve, thirteen, twenty, thirty*, plus a few subregular numerals, such as E 14–19. A few corresponding ordinal numerals are suppletive, e.g., E *first, second, third*, whereas *fifth* is subregular. The other ordinals are regularly formed from the right-hand head of the respective cardinal numeral via the suffix *-th* and its allomorph *-eth* (from *twenty, thirty*... ). The type frequency, i.e., the number of different heads is very small. Any quantitative measure, such as the ratio of potential words among conceivable words just applies to the rule of compounding heads, e.g.,

- (5) 79th, 89th, 99th = 9 → 9th + (70, 80, 90)

From the psycholinguistic point of view, storage of these few ordinal heads would be very easy and computing would only be necessary for the compounding of these stored heads with higher-number non-heads. Thus no rule of suffixing *-th* is needed for processing, and a fortiori no such productive rule is needed on psycholinguistic or computational grounds.

In our view, however, this instance is a very severe case of competence restriction, in the sense of restricting the potential domain (cf. Booij 1977; Bauer 1983, 88ff; Kastovsky 1982, 160–2; Rainer 1993, 117–24), but it is not a case of restricting productivity itself. In order to prove our point we will indulge in the Gedankenexperiment of inventing new non-natural numbers and of forming ordinals from them (following the type of E *a<sup>th</sup>, b<sup>th</sup>, ...*).

<sup>6</sup> This holds also for the related concept of rentability (Corbin 1987, 42–4; cf. the discussion in ten Hacken 1994, 177–8).

3.3. We start with German ordinal formation (irregular only with 1 and 3) which is done via two suffixes: *-te*, as in (6), *-ste* (homophonous with the superlative suffix), as in (7):

|     |             |          |         |         |        |       |              |         |
|-----|-------------|----------|---------|---------|--------|-------|--------------|---------|
| (6) | 2           | 3        | 4       | 10      | 11     | until | 19           | 0       |
|     | zwei/zwo    | drei     | vier    | zehn    | elf    |       | neun-zehn    | null    |
|     | zwei/zwo-te | drei-tte | vier-te | zehn-te | elf-te |       | neun-zehn-te | null-te |

|     |             |                     |  |       |             |             |                |
|-----|-------------|---------------------|--|-------|-------------|-------------|----------------|
| (7) | 20          | 21                  |  | until | 100         | 1.000       | 1.000.000 etc. |
|     | zwanzig     | ein-und-zwanzig     |  |       | hundert     | tausend     | Million        |
|     | zwanzig-ste | ein-und-zwanzig-ste |  |       | hundert-ste | tausend-ste | million-ste    |

*-ste* is analogously attached to:

|     |              |             |              |              |                  |
|-----|--------------|-------------|--------------|--------------|------------------|
| (8) | Milliarde    | Billion     | Billiarde    | Trillion     | Trilliarde...    |
|     | milliard-ste | billion-ste | billiard-ste | trillion-ste | trilliard-ste... |

Thus *-ste* is more frequent than *-te*, in terms of both type and token frequency.

The choice between the two suffixes depends on the right-hand head, as in

|     |                           |                     |
|-----|---------------------------|---------------------|
| (9) | 1001 = tausend eins       | 1003 = tausend drei |
|     | der/die/das tausend erste | tausend dritte      |
|     | 'the thousand first'      | 'thousand third'    |

But in (10) the numeral 1 has become part of a title, i.e., of a label or name.

|      |                                      |                                            |
|------|--------------------------------------|--------------------------------------------|
| (10) | 1001 nights = tausend-und-eine Nacht | 'thousand and one nights' (Arabian Nights) |
|------|--------------------------------------|--------------------------------------------|

When we refer to the last of these 1001 nights, we can form the ordinal

|      |                                              |
|------|----------------------------------------------|
| (11) | die tausend-und-ein-te/?erste/*ein-ste Nacht |
|------|----------------------------------------------|

This is clearly not a question of the combination with *und* 'and', as we can see from another Gedankenexperiment: In Mozart/Da Ponte's opera *Don Giovanni*, Leporello consoles his master's abandoned lover Donna Elvira with giving numbers of his previous lovers, ending with (in the German translation of Da Ponte's Italian text):

|      |                                                                   |
|------|-------------------------------------------------------------------|
| (12) | Aber/Ach in Spanien schon tausend und zwei, nein tausend und drei |
|      | 'But/Oh in Spain already 1000 and 2, no 1000 and 3'               |

Now in referring to Donna Elvira and to her immediate predecessor, he might have called them

(13) die tausend-und-dritte/\*drei-te/\*\*drei-ste; tausend-und-zwei-te/\*\*zwei-ste

and the next predecessor would be

(14) die tausend-und-erste/\*ein-te/\*\*ein-ste

because the base would be a real number and not a title (label, name). This case then gives a first indication that *-te* is the productive suffix, not *-ste*, and that if an ungrammatical derivation is formed, as in (14), it is still worse, if an unproductive suffix is used than a productive one.

The hypothesis of *-te* being the only productive suffix can be strengthened when looking at accepted ordinals derived from non-natural numbers and from 'how many' as in:

(15) x [iks] → x-te, n [ɛn] → n-te, wie viel? → wie-viel-te?

On the other hand we have the illegal poetic occasionalism (by Arno Schmidt):

(16) zum teufel-sten Male  
'for the devil-th time' (pejorative for 'nth')

But this illegal occasionalism must not count as evidence for productivity. In contrast, all ordinals formed from non-natural numbers, from invented numbers (taken from the Greek alphabet by a Gedankenexperiment) have *-te* suffixation, as in:

(17) a-te, b-te... , pi-te, phi-te, rho-te, tau-te, my-te, ny-te, ksi-te, chi-te, omikron-te, e/ypsilon-te...

whereas the examples with unstressed [a] in the second or third syllable, as in:

(18) ?alpha-te, beta-te, sigma-te, omega-te

sound funny to many informants (including university mathematicians). But this is a phonological, prosodic problem,<sup>7</sup> which cannot falsify the claim that *-te* suffixation is fully productive, whereas *-ste* suffixation is not.

**3.4.** Hungarian ordinal numeral formation is even simpler: after two suppletive forms we find two-/three-/four-way vowel harmony,<sup>8</sup> mobile vowels and vowel shortening, as in other domains of WF:

|      |          |           |            |             |          |
|------|----------|-----------|------------|-------------|----------|
| (19) | 1        | 2         | 3          | 4           | 5        |
|      | egy      | kettő/két | három      | négy        | öt       |
|      | első     | más-odik  | harm-adik  | negy-edik   | öt-ödik  |
|      | 6        | 7         | 8          | 9           | 10       |
|      | hat      | hét       | nyolc      | kilenc      | tíz      |
|      | hat-odik | het-edik  | nyolc-adik | kilenc-edik | tíz-edik |

Starting with 3, *-(V)dik* suffixation applies always to the right-hand head:

|      |                 |                  |                  |           |           |
|------|-----------------|------------------|------------------|-----------|-----------|
| (20) | 11              | 12               | 13               | 20        |           |
|      | tíz-en-egy      | tíz-en-kettő     | tíz-en-három     | húsz      |           |
|      | tíz-en-egy-edik | tíz-en-kett-edik | tíz-en-harm-adik | husz-adik |           |
|      | 30              | 40               | 60               | 100       | 1.000     |
|      | harminc         | negy-ven         | hat-van          | száz      | ezer      |
|      | harminc-adik    | negy-ven-edik    | hat-van-adik     | száz-adik | ezer-edik |

Later additions are (with weak suppletion between *-ó* and *-om*<sup>9</sup>):

|      |              |                    |                   |           |
|------|--------------|--------------------|-------------------|-----------|
| (21) | 1.000.000    | 1.000.000.000      | 1.000.000.000.000 | 0         |
|      | millió       | milliárd           | billió            | nulla     |
|      | milliom-odik | milliárd-(om)-odik | billiom-odik      | null-adik |

Again, the potential domain can be tested only, when new types of ordinal numbers are invented by a Gedankenexperiment, as from the non-natural numbers:

<sup>7</sup> Note the trochaic, morphotactically opaque variant *sieb-te* of the transparent, but dactylic ordinal *sieben-te* from *sieben* '7'.

<sup>8</sup> We cannot go here into the intricacies of choice between *-a-* [ɔ] and *-o-* [o] after velar vowels and between *-e-* [ɛ] and *-ö-* [ø] after labiopalatal vowels.

<sup>9</sup> Cf. the last-century poetic and dialectal variants *milliom/miljom* of *millió* 'million'.

- (22) m            n            x            a            b  
 [em]            [en]            [iks]            [a:]            [be:]  
 [emm-edik]    [enn-edik]    [iks-ɛdik]    [a:ɔdik]    [be:ɛdik]

or from Greek letter names, where there are no problems with:

- (23) [pi:]            [my:]            [ro:]            [ipsilon]            [omikron]  
 [pi:-edik]    [my:-edik]    [ro:-ɔdik]    [ipsilon-odik]<sup>10</sup>    [omikron-odik]

But phonological problems and uncertainties occur with bisyllabic Greek letters, such as (in a test):

- (24) [tɔu]                            [ɔlfɔ]  
 [tɔu-dik], [tɔu-ɔdik]    [ɔlfɔ-dik/ɔlf-ɔdik], rarely [ɔlfɔ-ɔdik]

But these phonological problems do not disprove productivity of Hungarian ordinal formation via the suffix *-(V)dik*.

**3.5.** In contrast, Russian ordinal formation via the suffix allomorphs (*-yj/-ij/-oj*) is not productive. One might claim to identify it, with interfixation of *-ev-* in two numbers that correspond to Hungarian (21)–(22), but no correspondents to the other Hungarian or German ordinals are possible:

- (25) nul'            [iks]            y            n            a            b            alfa  
 nul-evoj        iks-evyj        —        —        —        —        —

**3.6.** In Italian, ordinal formation via *-esimo* is totally productive, from 11 onwards:

- (26) **cardinal**        **ordinal**        80 (ottanta) +  
 1: uno            primo            ottant-un-esimo  
 2: due            secondo            ottanta-du-esimo  
 3: tre            terzo            ottanta-tre-esimo  
 4: quattro        quarto            ottanta-quattr-esimo  
 5: cinque        quinto            ottanta-cinqu-esimo  
 6: sei            sesto            ottanta-sei-esimo  
 7: sette        settimo            ottanta-sett-esimo  
 8: otto            ottavo            ottant-ott-esimo  
 9: nove            nono            ottanta-nov-esimo

<sup>10</sup> In a written ordinal formation test with 80 Budapest students (freshmen wanting to major in Hungarian language and literature) 44 produced *-odik*, 25 *-adik*.

|      |            |                 |             |          |             |         |
|------|------------|-----------------|-------------|----------|-------------|---------|
| (27) | 20         | 100             | 1.000       | n        | x           | a       |
|      | venti      | cento           | mille       | enne     | [iks]       | a       |
|      | vent-esimo | cent-esimo      | mill-esimo  | ennesimo | [iks-esimo] | a-esimo |
|      | [bi]       | ipsilon         | alfa        |          |             |         |
|      | [bi-esimo] | [ipsilon-esimo] | [alf-esimo] |          |             |         |

**3.7.** The morale of this section is: if an internal “lexical” constraint or the non-existence of appropriate/relevant loan words deprives a WFR of possible/adequate (new) bases, then this cannot be used as an argument for non-productivity. The analyst has to go beyond these obstacles by devising Gedankenexperimente and to show whether the potential domain of a WFR goes beyond its actual domain (cf. also Ladányi 2000).

If productivity is the ratio of potential (=legal) to conceivable words (**3.2**), then we can say, e.g., about the potential domain of non-natural numbers derived from the Roman alphabet that productivity of Hungarian, Italian and German ordinal formation is 100%, whereas Russian has just one legal ordinal (*iks-eyj*) among 31 conceivable ones, and thus has unproductive ordinal formation.

#### 4. German feminine motion

Pragmatically unusual situations may not only be created by Gedankenexperimente but actually occur under unusual circumstances. For illustration we want to discuss the case of sex of such animals whose sex is usually irrelevant for humans. Humans usually are not interested in the sex of animals whose sex makes no difference for them, e.g., the sex of bees is relevant economically, whereas the sex of other insects (wasp, hornet, flea...) is not. However, in a feuilleton (Standard 4.1.98) on the life of moles (G masc. *der Maulwurf*, umlaut pl. *Maulwürf-e*), a journalist dealt specifically also with the duties and habits of mole mothers after the birth of their children, and then used the non-existing fem. *Maulwurf-in*, not *Maul-würf-in* (with umlaut), although umlaut is regular in feminine motion of animals from masculine bases, particularly if they have an umlaut plural or other umlaut derivations (similar to humans):

- (28) Wolf ‘wolf’, pl. Wölf-e, adj. wölf-isch → Wölf-in; Storch ‘stork’, pl. Störch-e → Störch-in; Fuchs ‘fox’, pl. Füchs-e, adj. fuchs-isch → Füchs-in; Hund ‘dog’, pl. Hund-e, but adj. hünd-isch → Hünd-in; Affe ‘ape’, pl. Affe-n, but adj. äff-isch (vs. aff-ig) → Äff-in; Dachs ‘badger’ → Dächs-in.<sup>11</sup>



Here belong also the occasionalisms (cited by Wellmann 1975, 107):

- (29) die Taube 'dove' → Täub-in; die Katze 'cat' → Kätz-in

which are illegal, because derived from a base which is already feminine.

If there is no umlaut elsewhere, umlaut normally lacks in feminine motion as well:

- (30) Leopard 'leopard' → die Leopard-in; Schimpanse 'chimpanzee' → Schimpans-in; potential: Luchs 'lynx' → Luchs-in; Elefant 'elephant' → Elefant-in.

But umlaut in femine motion may occur even in such cases, as in:

- (31) Hase 'hare', pl. Hase-n → Häs-in; fem. Ratte 'rat' → (illegal) occasionalism Rätt-in (G. Grass)

Thus we may posit the rule of feminine motion of animal names:

- (32) If an animal name is masculine and has an umlautable root, then the feminine (denoting the respective female animal) is formed via the suffix *-in* and umlaut of the umlautable root vowel.

Human correspondents of (28) are (cf. Doleschal 1992):

- (33) Gott 'god' → Gött-in; Arzt 'physician' → Ärzt-in; Papst 'pope' → Pöpst-in Johanna; Graf 'count' → Gräf-in; Bauer 'peasant' → Bäuer-in; Bub(e) '(bad) boy' → Büb-in; Koch 'cook' → Köch-in; Rat 'councillor' → Rät-in; Tor 'fool' → Tör-in; Narr 'fool' → Närr-in; Sachse 'Saxon' → Sächs-in; Franke 'Frank' → Fränk-in; Franzose 'French' → Französ-in; Jude 'Jew' → Jüd-in; Lands-mann 'compatriot' (and other compounds with the same head) → Lands-männ-in.

Exceptions, with umlaut elsewhere, but not in feminine motion, are:

- (34) Herzog 'duke' → Herzog-in; Flame 'Flemish' → Flam-in (?Fläm-in); Lappe 'Sami' → Lapp-in.

Thus *Maulwurf*, with its plural *Maulwürfe* falls into the potential domain of umlauting feminine motion. Thus the above-mentioned occasionalism *Maulwurf-in* contradicts this WFR. However, this seems to be an error of the journalist, in

<sup>11</sup> Cf. the diminutive *Dächs-el*. In a test, some informants produced *Dachs-in*.

view of the poetic occasionalisms cited by Wellmann (1975, 492, notes 60, 64), maybe unknown to this journalist:

- (35) Maulwürf-in; Frosch 'frog' → Frösch-in; Kauz 'screech-owl' → Käuz-in (all Th. Mann); Schwan 'swan' → Schwän-in (E. Strittmatter); fem. die Maus 'mouse' → (illegal) Mäus-in (Th. Mann, cf. (47); (49)); personalised Glockenton 'sound of a bell' → Glockentön-in (Chr. Morgenstern); Kopf 'head' → Köpf-in (A. Schmidt: illegal, because inanimate)

If the conditions of (32) apply, then there is, as expected, no umlaut in such poetic neologisms:

- (36) Mond 'moon' → Mond-in (E. Penzoldt); Uhu 'eagle owl' → Uhu-in (Chr. Morgenstern); Kondor 'condor' → Kondor-in (A. Schmidt); die Wolke 'cloud' → (illegal) Wolk-in (A. Schmidt)

And similar to excessive umlaut in (31), and in contrast to lack of umlaut with humans in (34), there is excessive umlaut in the poetic occasionalisms:

- (37) Spatz 'sparrow', pl. Spatz-en → Spätz-in (Chr. Morgenstern); Rabe 'raven' → Rüb-in (Th. Mann).

If we extend empirical testing to a Gedankenexperiment, where we imagine a potential world where it would be necessary to apply feminine motion to other animals, we obtain (from all our informants so far), as predicted:

- (38) der Floh 'flea', pl. Flöh-e → Flöh-in<sup>12</sup>

and a confirmation of the occasionalisms in (35).

In conclusion we may state: since the type frequency of feminized animal names is small and their distributional patterning not totally identical with their human correspondents, it may seem adventurous to postulate a productive WFR on the level of the potential system. Moreover, we cannot apply the productivity criteria of section 5, i.e., apply the rule to recent but sufficiently familiar foreign animal names or to abbreviations, and there are severe pragmatic restrictions on forming new ones. Still we hope to have been able

<sup>12</sup> Erben (1993, 46) forms *Flöh-in*, but rejects it for pragmatic reasons. However, already the text worlds of Goethe's (Faust I, and Mussorgski's) "Flohlied" and E. T. A. Hoffmann's "Meister Floh" provide a pragmatic base for feminine motion.

to outmanoeuvre these obstacles and to test the grammatical productivity of this postulated WFR.

## 5. Gradation of grammatical productivity

Our examples so far have been fully compatible with our thesis that WF productivity, as grammatical, rule-governed creativity, is a primitive and prototypical property of WFRs within the potential system of grammar, in analogy to grammatical rules of inflection, of syntax, and of phonology. This implies that unproductive rules are marginal within morphological grammar (including WF). This is clearly a qualitative, not a quantitative conception of productivity. Therefore, if we assume that, as many other concepts of naturalness theory (our underlying framework), also productivity is gradual, then also this gradualness must be of a qualitative, grammatical nature. Our proposal for establishing a hierarchy of grammatical productivity criteria of WFRs is based on an analogical hierarchy established for inflectional morphology (in Dressler 1997; Dressler–Thornton 1996). Some justification for the following hierarchy will be given immediately, the theoretical bases will be further elaborated in section 6.

(a) The highest degree of derivational productivity is obtained with Wurzel's (1984) secondary productivity, i.e., when (i) even new foreign words are integrated (sc. we start from the general assumption that it is more difficult to integrate words coming from a foreign language [marked alternative] than indigenous words [unmarked alternative]), (ii) when these foreign words have unfitting properties and (iii) when these unfitting properties are "fitted", i.e., accommodated to the properties of a derivational class or of the respective language-specific system adequacy in particular. Clearly a WFR must have optimal productivity in order to overcome the difficulties of (i) and (ii) and to enforce accommodation (iii).

Accommodation of unfitting phonological shapes can be exemplified with F *chaussée* 'road' loaned into Russian as [š'o'se], which is phonologically non-integrated because of unstressed [o] and stressed word-final [e], and morphologically non-integrated, because it is indeclinable, thus this loanword has three unfitting properties. The derived adjective [ša'sej]-*nyj* and the derived verb [šas]-*irov-at'*, however, are both phonologically and morphologically integrated: the unstressed vowel is centralised, the final vowel is changed into a diphthong in the adjective and deleted in the verb, and the derivational suffixes allow inflection.

According to Wüster's (1979) definition, 'guest words' are phonologically and phonetically non-integrated words of one language when cited in another language. We assume that morphological integration presupposes phonological and phonetic integration. Therefore we claim that no guest words, i.e., words with conscious approximation of original foreign pronunciation, accept indigenous affixation. An example is G *Komputer-isier-ung* 'computer-iz-ation': if the first vowel is pronounced in the English way as an a-schwa and if the three voiceless stops are aspirated in the English way, then, at least in Austrian German, the above German derivation is impossible. Another German example (supplied by W. U. Wurzel, personal communication) is G *Szient-o-loge* [stsiento'lo:gə] 'scient-o-logist': if the first syllable is pronounced in the English way ([sai]), then the German derivational suffix may not replace the English one.

Similarly in the Hungarian adaptation of E *to print* as *print-el* 'prints', the lexical root may not be pronounced in the English way (aspiration of stops, rhotic approximant), or in *deep-jumping-ol*, the first [i:] must not be diphthongized, and (ng) must be pronounced as two consonants (nasal plus voiced velar stop) instead of velar nasal only (as in English), if the morphological adaptation via the suffix *-ol-* (section 2.1) is to be applied.<sup>13</sup>

A special case of criterion (a) is represented by obligatory derivational adaptation, as in Hungarian obligatory adaptation of foreign verbs via the denominal verb formation suffix of postvocalic *-l*, postconsonantal *-e/o/öl*:

(39) *print-el*, *menedzs-el*, *szév-el* ('saves'), *klikk-el* ('clicks')

Another type of morphotactic adaptation can be identified in the adaptation of Romance thematic vowels into the German athematic verb system via substituting the Romance thematic vowel with the German verb formation suffix *-ier-* (since the late medieval period):

(40) F *march-er*, *ras-er* > G *marsch-ier-en* 'march', *ras-ier-en* 'shave';  
 LL *salv-a-re* > *salv-ier-en* 'save';  
 It *collaud-a-re* > Austrian G *kollaud-ier-en* 'ratify the construction of a building'

<sup>13</sup> Moravcsik (1975; cf. 1978) claims that these English verbs are loaned as nouns into Hungarian and then verbalized again. This entails the assumption of systematically not existing abstract nouns. Moravcsik's more general claim that verbs are never loaned directly is especially problematic for Semitic languages which abundantly loan foreign verbs into their verbal systems.

(b) A hierarchically lower criterion of productivity is represented by Wurzel's (1984) primary productivity in the integration of (new) loan words, because here the class-defining properties of the loan word already fit into the receiving class, i.e., only the first difficulty (ai of criterion a) has to be overcome. Examples are the German denominal adjectives, derived from French or English loan words:

(41) gelee-ig, rass-ig, trend-ig, zwei-etag-ig,<sup>14</sup> rass-isch, film-isch; sport-lich

or the German derivations from English athematic verbs, loaned as German athematic weak verbs:

(42) an-klick-en, aus-klick-en < click, out-sourc-en, lift-en (sc. 'a face'), etc.

In the following Russian example, adjective derivation from a loaned noun triggers the normal morphonological rule of palatalization:

(43) E bulldog > R buldog → adj. buldož-nyj

Hungarian denominal adjective formation with the suffix *-s* (*-a/o/e/ös*) includes all the meanings that it has in derivations from indigenous words, i.e., characterizing a person, an object or an activity (belonging to something or dealing with the given activity):

(44) multimédiá-s, marketing-es, graffiti-s, jet-ski-s, internet-es, intervenció-s, privatizáció-s, koncesszió-s

or meaning 'similar to, just like something or somebody':

(45) raggie-s, blues-os, kolakowski-s

or, in relation to illness, referring to the state of suffering from that illness:

(46) AIDS-es [e:ts:ɛʃ], hepatitis-B-s [hepətisbe:ʃ]

or, with names of professions, characterizing the person's behaviour:

<sup>14</sup> Note that adjectival *-ig* suffixation is only productive with noun bases but unproductive with verb bases. Thus the criterion of productivity supports Aronoff's (1976, 47) Unitary Base Constraint in this and many other cases.

- (47) bróker-es, menedzser-es ‘broker/manager-like’

The two already-mentioned Hungarian denominal verb-formation suffixes *-(o/e/ö)z-* and *-(o/e/ö)l* apply to foreign loan words with fitting criteria, as in:

- (48) *bridzs-el* = *bridzs-ez-ik* ‘(s)he is playing bridge’

The integration of bases imported from a foreign lexical system into the indigenous WFR component thus presupposes a high degree of grammatical WF productivity. In contrast to Wurzel, we have argued that his “secondary productivity” is of greater importance than his “primary productivity”. Of course, if there are no loan words with unfitting properties, which are potential bases of a WFR, i.e., if all incoming loan words have fitting properties, we cannot be sure whether this WFR has type-a productivity or just type-b productivity.

(c) Hierarchically still lower as a productivity criterion is derivation from indigenous abbreviations. Note that abbreviations are not formed via grammatical WFRs but by extragrammatical means (cf. Dressler–Merlini Barbaresi 1994, 36–40) and are therefore both marginal and marked within the lexical stock of a language. We rank them lower than foreign words, because they appear to be less foreign to the indigenous lexicon than truly foreign loan words. Examples of denominal nouns are:

- (49) *G SPÖ-ler* [ɛspe:ølər], *ÖVP-ler*, *CDU-ler* ‘member of the Austrian Sociodemocratic/People’s Party, of the (German) Christian-Democratic Party’;  
*R emgeuś-nik* ‘student of the MGU (Moscow State University)’, *vuz-ov-ec* ‘student of a college’ (base *VUZ* with interfix *-ov-*)

Compare Hungarian denominal adjective formation with the suffix *-s* (*-a/o/e/ös*) with the meaning of ‘belonging to (an organization)’:

- (50) *MSZP-s* [ɛmɛspe:š], *fidesz-es*, *MDF-es* [ɛmde:ɛfɛš]; *kft-s* [kæfte:š] ‘being a party member of the parties MSZP, Fidesz, MDF; being/working in a limited liability company’

or denominal verb formation with the suffix *-z-*, with the meaning of either ‘to take part in the activities of an organization’ or ‘to speak pejoratively about an organization’:

- (51) *kft-z-ik*; *MSZP-z-ik*, *fidesz-ez-ik*, *MDF-ez-ik* ‘(s)he works in a limited liability company; works in/for the M/F/M party’

Or take abbreviations of primary- and secondary-school students:

- (52) matematika > matek → matek-oz-ik 'mathematics - maths - (s)he does maths'

Now we seem to have a problem for the criteria hierarchy, that does not exist in inflectional morphology: the synonymous *-(o/e/ö)l* suffixation applies to foreign loan words (criterion a, b), but not to abbreviations (criterion c), which appears to contradict the hierarchy. Of course, one could counter that *-(o/e/ö)l* suffixation applies only to monosyllabic bases, whereas all of these abbreviations are polysyllabic. However, if one tries to form potential, actually not existing verbs from monosyllabic (even non-existing) abbreviations, then still only the suffix *-z-* seems to be possible, as in:

- (53) Magyar Anyák/Építészek Szövetsége > MASZ, MÉSZ → MASZ-oz-ik, MÉSZ-ez-ik  
'(s)he works in/for the alliance of Hungarian mothers/architects'

An explanation for this apparent contradiction to our hierarchy is that abbreviations constitute a marginal, rather small and homogeneous set of bases which appear not to allow productive WFR competition.

(d) Hierarchically still lower as a productivity criterion is the shift of a derivative from one (prototypically) recessive to another (prototypically more stable and thus more productive) WFR within a derivational microclass. Due to the low hierarchical rank of this criterion, productivity of the target WFR may be very slight. German examples of such suffix replacement are (see Pounder 1987; Osman 1994):

- (54) G-ig > -isch, when pejorative: argwöhn-ig > argwöhn-isch 'distrustful' (← Argwohn),  
neid-ig > neid-isch 'envious' (← Neid)

In German deadjectival quality nouns the suffix *-e* (often homophonous with the fem. adj.) has been substituted with *-heit/(-ig-)keit*, as in (see Osman 1994):

- (55) Fein-e > Fein-heit 'fineness', Gleich-e > Gleich-heit 'equality' (but non-homophonous compounds Dach/Keller-gleich-e), Heiter-e > Heiter-keit 'cheerfulness', Klär-e > Klar-heit 'clearness', Klein-e > Klein-heit 'smallness', Leicht-e > Leicht-igkeit 'easiness', Rau-e > Rau-heit 'harshness', Schön-e > Schön-heit 'beauty', Süß-e > Süß-igkeit 'sweetness', Trocken-e > Trocken-heit 'dryness', Wohlfeil-e > Wohlfeil-heit 'cheapness'

Rather recent is the shift from the complex Hungarian suffix *-íroz-*, used to form verbs from foreign names and nouns, to the normal, simple indigenous suffixes *-(e/o/ö)z-* and *-(e/o/ö)l* as in:

- (56) vagon-íroz > vagon-oz 'loads into a train-carriage (cf. G ein/aus-waggon-ieren)', park-íroz > park-ol 'parks', patent-íroz (< G patent-ier-en) > patent-ol, szald-íroz (< G sald-ier-en) > szaldó-z (with the N szaldó < It, G Saldo), storn-íroz (< G storn-ier-en, with Austrian colloquial pronunciation [št-]) > sztornó-z (with the N sztornó < It, G Storno)

(e) The last and hierarchically lowest criterion is WF productivity of affixations, as directly observable in fully grammatical indigenous neologisms formed from indigenous bases. This is productive derivation from unmarked bases, sc. from fully integrated indigenous bases. This criterion can be further subdivided into the following subhierarchy (e1)–(e9):

(e1) The highest subcriterion appears to be rule application in derivations from non-actual bases, i.e., via intermediate false steps (Rainer 1997a), as in the following Russian examples (taken from Zemskaja 1996, 108f):

- (57) R tret'emir-izacija 'process of transforming a country of the second world (such as Russia) to a country of the third world' (← V tret'emir-iz-at') ← tret'ij mir 'third world', velosiped-izacija 'introduction of bicycle zones/ways' (← V velosiped-iz-at') ← velosiped 'bicycle', bandit-izacija 'process of a locality/organisation getting infested with criminals' (← V bandit-iz-at') ← bandit 'criminal'

It is not clear to us whether in the formation of (58) there has been an intermediate step of an actual or only of a potential verb.

- (58) E finland-is-ation, G Finnland-isier-ung ← Fin(n)land

(e2) The potentiality of a derivation may become actual in colloquial, non-standard oral speech, i.e., such neologisms are excluded by norms of the standard, but are still legal and thus potential words of the standard. Zemskaja, Kitajgorodskaja and Širjaev (1981, 72) give many different types of such "neuzual'nije slova" ('non-usual words'), which may be quantitatively more important in recent Russian than in other languages due to hidden private life in Soviet times. A few examples<sup>15</sup> are:

- (59) kastrjul'eč-nost' 'a situation where pots are everywhere in the kitchen' < kastrjul'ja 'pot'

<sup>15</sup> For Hebrew examples see Berman (1987).



- (60) žolt-ost' 'yellowness' in the sentence (*ibid.* 86):  
 Kakoj u tebjā halat žoltyj! 'What a yellow gown you have!' – Ja jego kupila za etu žolt-ost' (= za to, što on žoltyj) 'I have bought it because of its yellowness (= because it is yellow).'

Educated Russian speakers, due to their restrictive normative attitude against colloquialisms, strongly tend to evaluate such colloquial neologisms as unacceptable, although they agree with their morphological legality (within the potential system).

(e3) In case of productive rule competition, we often find just one actual word formed via one of the productive competing WFRs instead of several potential variants (formed from the same base via the other competing WFRs). This then appears to be a case of lexical blocking (synonymy blocking). However, such potential variants may be formed under special conditions, such as in poetic play with words, e.g., by the German theatre critic Alfred Kerr (see Erben 1996; 1997):

|      |                   |                         |                                  |
|------|-------------------|-------------------------|----------------------------------|
| (61) | base              | accepted                | potential/legal ad hoc formation |
|      | V herum-kriech-en | N Herum-kriech-erei     | Herum-ge-kriech 'creeping'       |
|      | N Skizze          | Adj skizz-en-haft       | skizz-ig 'sketch-y'              |
|      | N Episode         | Adj episod-isch/-enhaft | episod-ig 'episod-ic'            |
|      | N Nachdruck       | Adj nachdrück-lich      | nachdruck-sam 'emphat-ic'        |

Such examples, as well as literary manuscript variants (cf. Pilhak 1975 on adjective formation in Georg Trakl's poems) show that lexical blocking refers to accepted lexical usage on the level of language as norm but not to the potentiality of productive WFRs on the level of competence, as has been stated already by Rainer (1993, 115f). This is similar to potential variation in inflectional morphology (cf. Dressler 1997; Dressler-Dziubalska-Kořaczyk 1997), another argument for the basic unity of inflectional and derivational morphology (thus against the separation thesis). Further evidence for the assignment of synonymy blocking to language as institution is its dependence on acceptance by relevant segments of the language community. This does not only hold for terminology within the community of experts but also for the community of opinion leaders and journalists in the mass media (cf. Alloni-Feinberg 1977).

(e4) The last type of potential but not accepted words is represented by ad hoc formations that have not yet been accepted, although they are completely legal, such as:

- (62) G bier-ig ← Bier 'beer': Ein bier-ig-es Gasthaus/Fest 'a beer-y inn/feast' (a 1998 announcement in Salzburg/Vienna)

- (63) H *kóstol-ás-os* (teszt) 'trying-out (test)', adj. from *kóstol-ás*, existing action noun of V *kóstol* 'to try out' (1998 ad in Budapest); el-unalmas-odik 'PFV+ boring+ deadjectival verbalizing suffix = to become boring' (conversation about a movie)

Less important subcriteria are offered by accepted neologisms (e5)–(e9):

(e5) The most innovative ones are those formed from neologisms which are morphologically complex, i.e., are neologistic derived bases. Thus any new noun in E *-ism*, G *-ismus*, H *-izmus*, etc. can be derived to E N/adj. *-ist*, H N/adj. *-ista*, G N *-ist*, adj. *-ist-isch*, etc. Or from adapted foreign nouns referring to types of creams, perfumes, etc., the following Hungarian verbs (with the meaning of 'putting that type of cream, perfume, etc. on oneself/somebody's face') are formed with the suffix *-(o/e/ö)z*. From such neologisms, passive reflexives (usually also with the prefix *be-*, as above in (e1)) can be derived via suffixation of *-ó/őd-ik-*:

- (64) *be-dezodor-oz-ód-ik*, *be-niveá-z-ód-ik*, *be-tonik-oz-ód-ik*, *be-oxy-z-ód-ik*, *be-givenchy-z-ód-ik*

Analogously, deverbal reflexive active formation with *-ko/e/ödik* can be applied to the same bases, because they are Hungarian verbs, indirectly derived from a loan-noun (and not from a loaned verb), e.g.,

- (65) E tonic > H N *tonik* → V *tonik-oz* → act. refl. V *tonik-oz-kodik*

(e6) refers to less complex neologisms, i.e., those derived from indigenous neologistic simplex bases, e.g.,

- (66) G *Kitsch* 'trash' → adj. *kitsch-ig*, F *gas* → *gas-eux*

(e7) Still less innovative and therefore providing less evidence for high productivity are recent neologisms from long-accepted, traditional words, such as the German pedagogical terms (denominal adjectives):

- (67) *kind-haft*, *familien-haft* ← Kind 'child', Familie 'family'

They may fill accidental lexical gaps (cf. Corbin 1997a) or be formed because of some other pragmatic reasons. The formation of many of them is stimulated by a foreign term, being structurally independent, such as:

- (68) F milieu → G Um-welt (lit. 'ambi-world'); E drug-addict → H drog-os (lit. 'drugg-ish'); F bicyclette 'bicycle' → coll. Bret marc'h-houarn 'horse (of) iron', coll. G Draht-esel 'wire-donkey'

(e8) Even less innovative are loan translations, because the choice of basis and WFR is stimulated by the foreign source, as in:

- (69) coll. G Draht-esel → coll. H drót-szamár 'wire-donkey = bicycle'; E brows-er → H böngész-ő; R luno-xod (lit. 'moon-go(-er)' = 'moon vehicle') → H hold-jár-ó (lit. 'moon-walk-er')

cf. the following Latin words formed after Greek models (André 1971, 31, 18):

- (70) Lat. suffix -men(tum) = Gk. suffix -ma: cogita-men < nóe-ma 'thought', nova-men < néo-ma 'innovation'; abomina-mentum < bdélyg-ma 'abomination'.

(e9) Still less conclusive are those rule-governed neologisms whose creation has been stimulated or at least facilitated by analogy (cf. G *ent-drei-t* in 2.4 (a)). One instance is:

- (71) H mars-jár-ó (lit. 'Mars-walk-er'), formed after H hold-jár-ó (lit. 'moon-walk-er' = 'moon vehicle') in (69).

## 6. Bases and consequences of the productivity hierarchy

Criteria for measuring degree of productivity must be tested with performance data, be it of off-line tests with informants (i.e., testing competent native speakers in their performance of either production or evaluation), or by on-line tests (which we were unable to do), or by searching through corpora (the results of implicit off-line tests, so to say) in a quantitative and qualitative way (for the latter cf. section 10). This is language processing evidence. The theoretical basis for tying degree of productivity to processing evidence is the assumption that the more productive a mechanism is, the easier it can be used in production, reception, evaluation (cf. Frauenfelder-Schreuder 1992; Baayen 1994).

Our elaboration of the bases of the hierarchy of criteria for grammatical productivity must start with the concept of gradualness. Gradualness, instead of discreteness, is an essential property of any model which works with prototypes, Natural Morphology being one of them (cf. Dressler 1989; Taylor 1991).

We have assumed in 2.1 that productivity is a prototypical property of rules, such as in inflection, syntax and phonology. This fits to the function of WFRs to form potential words whose number is in principle non-finite. That is, the actualization of rule productivity can be restricted in the rule domain (cf. 3) or for pragmatic reasons (cf. 2.5), but the mechanism of rules within their domain is basically (or, in terms of a flow chart, initially) unrestricted, thus completely productive.

Diachronic grammatical change, however, may lead to either loss of productivity or origination of productivity and, between these opposed poles of full productivity and unproductivity to decrease or increase of productivity (cf. van Marle 1988; Sánchez Miret et al. 1997). This is theoretically deducible from the assumption of gradualness, and empirically supported by diachronic and acquisitional evidence that productivity is not an all-or-nothing notion.

Now we propose that the ensuing synchronic gradation of grammatical productivity should be measured according to the relative size or severeness of structural obstacles a rule has to overcome in order to realize or perpetrate its productivity and that the criteria for measuring productivity should be hierarchically ordered according to the same gradation. Pragmatic obstacles are a less reliable basis of measurement, because they are of an entirely different kind than structural properties of grammar (including the lexicon). The fundamental grammatical property of a structural unit is its (degree of) language-specific system adequacy.<sup>16</sup>

Old indigenous words are clearly more system-adequate than foreign words entering the target language. First, they are traditional and thus more familiar on the level of language as a social institution, while new<sup>17</sup> words do not belong to the core of the lexical system of a language. At least, within the primary lexicon of simplex words, stored lexical items form the core of the potential lexicon. Second, newly loaned words are easily recognized as foreign, in contrast to new indigenous words. As a consequence, we rank the application

<sup>16</sup> The domain of the third subtheory of Natural Morphology, as pioneered by Wurzel (1984), cf. Dressler (1997).

<sup>17</sup> There is a methodological problem here in delimiting what a new word is: (i) Individual speakers know only part of the word stock: e.g., the German term *Lern-er* 'learn-er' had been created several times before it became a known accepted neologism, due to the model of E *learn-er*. (ii) The subjective impression of novelty depends on many factors, cf. Barz (1998).

of rules to foreign words<sup>18</sup> as more convincing evidence for rule productivity than their application to indigenous words (criteria (a), (b) vs. the rest), and the application to new indigenous words higher than the application to old ones (criteria (c), (e1)–(e6) vs. (e7)–(e8)).

Among new foreign words, those with non-fitting structural properties are clearly more probative than those with fitting properties, because they present a greater structural obstacle for rule application (criterion (a) vs. (b)).

Abbreviations are usually also new words (rather few go back to the 19th century, most are very recent). Moreover, since they are not formed via grammatical WFRs but by extragrammatical means, they are marginal within the lexical system of a language, i.e., they are less system-adequate than the rest of indigenous words (criterion (c) vs. (d), (e)). We rank abbreviations lower than foreign words, because they appear to be less foreign to the indigenous lexicon than truly foreign loan words (thus criterion (c) after (a), (b)).

Potential words are more probative bases than actual words (criterion (e1) vs. (e2)–(e8)), cf. above.

Criteria (e2) and (e3) are off-shoots of the potentiality status of (e1), so to say on side-branches of the hierarchy: colloquial realization of potential but not actual words of the standard (criterion (e2)) involves a style-shift and a distance between colloquial and standard language which faintly recalls the distance between foreign and indigenous words. Thus realization of potential words in colloquial usage only is more probatory of productivity than realization in standard language itself. Unfortunately, in empirical research, this difference overlaps with the distinction between written and oral language (cf. ongoing research by Plag et al. (2000) on greater productivity of WFRs in written than spoken English).

Criterion (e3) (potential alternatives) points to the additional, independent obstacle of lexical (synonymous) blocking which may obstruct the actualization of a potential word. Thus in empirical counts of new derived words this additional factor must be taken into account.

Not-accepted but legal new words (criterion (e4)) are more valuable for measuring (degree of) productivity than accepted ones, because there is less interference of institutional norms with the potentiality of the language system.

<sup>18</sup> In cross-linguistic comparisons, there is the methodological problem that some languages (e.g., Italian, cf. Petralli 1996, 23–4) loan foreign words much easier than others (e.g., Icelandic).

Among accepted new words those derived from neologistic bases (criteria (e5)–(e6)) are more valuable for measuring productivity than those derived from old words (criterion (e7)), because rule application has to overcome the factor of newness, appealed to above, and among neologisms, complex bases (criterion (e5)) are more valuable than simplex bases (criterion (e6)), because rule application is more complex in language processing.

Neologisms which are the effects of loan-translation (criterion (e8)) or analogy (criterion (e9)) involve a factor favouring application of a rule, instead of a disfavouring obstacle, and are thus less valuable for measuring productivity.

This leaves us with the problem where to insert criterion (d) into the hierarchy of criteria. If a derivative replaces suffix *x* with suffix *y*, then (*ceteris paribus*) suffixation of *y* should be more productive than suffixation with *x*, because synchronically a more productive rule takes precedence over a less productive one.<sup>19</sup> Such diachronic change, if relatively recent, may then be taken as the result of a still existing synchronic gradation of productivity. This is only a safe criterion if such affix replacement is not isolated but a recurring phenomenon. Next, we assume that there is a greater obstacle to replacing an existing word than to form a new word. Therefore this criterion (d) should be considered to be the most important criterion of all. However, this criterion has to do with indigenous words (including well-integrated loan words and, possibly, abbreviations), and therefore it should be ranked after criterion (c). Moreover analogical influence is a factor favouring replacement, which downgrades the value of this criterion. Unfortunately well-studied cases of such recurring affix replacements are too rare for deciding this issue.

With these caveats in mind, we may use these five criteria (plus subcriteria) for establishing degrees of grammatical productivity of WFRs, in strict parallel to productivity of inflectional morphology. In addition, restrictions of rule domain (see 3) and pragmatic factors must be taken into account for relating degree of productivity to type and token frequency and for interpreting psycholinguistic tests.

One important theoretical issue not discussed here is the interaction of productivity with morphosemantic transparency (cf. van Marle 1988). As stated elsewhere (Dressler 1997, 11), both properties have to be clearly distinguished, although they are connected in their metatheoretical foundation, and although morphosemantic transparency may favour the application of morphological

<sup>19</sup> One exception is resegmentation, where relative morphotactic transparency appears to play a bigger role than relative productivity (cf. Fuhrhop's 1998 third chapter "Suffixreanalyse").

rules in processing, which may have consequences in diachronic change (cf. van Marle 1988).

## 7. Productivity in a Stratified Lexicon

The word stock of many languages is stratified, usually according to a feature [ $\pm$ foreign], and this may extend to WF (cf. Rainer 1993, 129f). Thus suffixation with E *-ity*, G *-itüt*, H *-itás* is only productive with Latinate bases, whereas suffixation with E *-ness*, G *-h/keit*, H *-sá/ég* is (potentially) productive with all adjective bases, although actually more productive with native than with Latinate bases, due to rule competition with E *-ity*, G *-itüt*, H *-itás* in the Latinate lexical domain. If we concentrate on affixation which is restricted to the [+foreign] domain, we must reduce the hierarchical five-point scale of section 5 to the following two-point scale:

(a) fitting of unfitting properties:

The following type of English loan-adjectives in Hungarian computerese must adapt and enlarge the shape of their (pseudo)suffix with *-is*:

- (72) E virtu=al > H virtu=ális, compat=ible > kompat=ibilis, digit-al > digit=ális, cf. older profit-able (or rather G profit-abel < F profit-able) > profit=ábilis, structural > strukturális (N struktúra), cultur-al > kulturális (N kultúra), experiment-al > experiment=ális (puristic indigenous N kísérlet).

Latin adjectives in *-alis* and *-bilis* had been taken over in this (orthographic) base-form, e.g.,

- (73) liberális, rivális, duális, spektábilis

and even recent English loans have been adapted to this Latinate stock of Hungarian. Since, however, such Latinate adjective formation suffixes are never suffixed to non-Latinate words in Hungarian itself, adaptation (fitting) is evidence for system adequacy, stability, and productivity within the small Latinate stock of Hungarian.

Similarly, Latin adjectives in *-icus* were simply taken over into Hungarian, such as:

- (74) la=ikus, trag-ikus, aul=ikus

Accordingly, Latinate adjectives with the suffixes E *-ic*, F *-ique*, It *-ico*, G *-isch* have these suffixes substituted with H *-ikus*, when integrated into Hungarian, as in:

- (75) G N Kult(-us) < Lat cultus, adj. kult-isch > H N kultusz, adj. kult-ikus; G Fakt < Lat factum, adj. fakt-isch > H occasionalism (1997) adj. fakt-ikus

But, again, *-ikus* is never suffixed to indigenous Hungarian bases. This distinguishes these cases from the fitting of foreign verbs via the suffix *-(e/o/ö)l-* in section 5 (a), because this suffixation also verbalises indigenous Hungarian nouns.

Therefore the above adaptations of adjectives represent an accommodation which is rather comparable to the fitting of foreign verbs into Hungarian ones via suffixation of *-ál*, which never applies to indigenous Hungarian bases:

- (76) E to edit, install, adapt > H edit-ál, install-ál, adapt-ál

Diachronically, though, this suffix split from the before-mentioned postvocalic allomorph of the suffix *-(e/o/ö)l-*, added to Latin, Italian and Slavic thematic vowels /a/, with the typical vowel lengthening of final ⟨a⟩ to ⟨ál⟩. Thus only H *edit-ál* is a clear example of the suffix *-ál*, whereas the other verbs of (76) may be related to Neolatin or Italian verbs in *-are*.

(b) Examples of Latinate-only suffixation to bases with fitting properties, i.e., without any further accommodation, abound. Relevant nominalising suffixes are, e.g.,

- (77) deverbal G -ation = H -áció, deadjectival G -itāt = H -itás.

When we combine the results of sections 5 and 7, we may confidently state that WFRs which are productive both in indigenous and “foreign” (e.g., Latinate) WF, are more productive than those WFRs which are productive only in one of these two strata.

## 8. Hierarchical paradigmatic organization of WF

As has been abundantly shown by van Marle (1985; cf. Pounder 1987; Bauer 1997), the syntagmatic relations between base and derivation have to be supplemented with the paradigmatic relations among different derivations. What



we claim and want to demonstrate here, is that these paradigmatic relations are hierarchically organized, in a way which recalls the hierarchical organization of inflectional paradigms (cf. Dressler–Thornton 1996; Dressler–Dziubalska-Kolaczyk 1997; Pöchtrager et al. 1998).

We define a **derivational paradigm** as a set consisting of a base and of its parallel derivations which belong to the same category and share the same or nearly the same word-formation meaning. Parallel sets form a **derivational microclass**, e.g., the microclass of adjectival relational derivations from nouns, as in G *Kind* ‘child’ → adj. *kind-lich* and pejorative *kind-isch*, *Weib* ‘wife’ → adj. *weib-lich* and pej. *weib-isch*, etc. (see Pounder 1987). Also here we have to differentiate (cf. Zemskaja 1992, 17–9) actual, concrete WF paradigms, such as the above, and the corresponding abstract WF paradigm within language as a potential system, which is the expression of the word formation meanings of the respective WFRs as applied to a specific base (such as G *Kind*). The derivational microclass is then established by the systematically recurring complementary distribution of the WF meanings of the WFRs taking part in the abstract paradigms.

In our above example, there is a systematic connotational meaning difference between the two WFRs involved. Another type of microclass is established by WFRs with synonymous WF meanings, as in Hungarian frequentatives:

- (78) prod. -ga/et- and unprod. -o/e/öz-: küld ‘(s)he sends’ → küld-öget = küld-öz ‘sends repeatedly’, hord ‘carries’ → hord-ogát = hord-oz ‘carries repeatedly’, told ‘lengthens’ → told-ogát = told-oz ‘lengthens repeatedly’

The concept of a derivational microclass thus represents a meaningful generalization over (sometimes only partial) systematic, grammaticalized WF synonymies within the lexicon of a language. Beard (1995, 155, 175, 202) differentiates between sets of grammatical functions or functional derivations on the level of meaning (morphosemantics) and sets of rules and classes of derivations on the level of expression. Our claim is that within the potential domain of WF such sets generally and prototypically coincide, which allows the establishment of microclasses. In the case of unproductive rules, however, the domain of an unproductive rule may not coincide with the morphosemantic domain of the microclass. Such unsystematic non-coincidence results in **isolated derivational paradigms**, i.e., paradigms whose WF set of members does not correspond to that of any other derivational paradigm, as in:

- (79) H lép ‘goes one step’ → lép-eget = lép-ked = lép-del = lép-des = lép-degél ‘goes step by step’

where only the first member is derived via a productive WFR, and where no other WF paradigm contains the application of all of these five WFRs. Thus Beard's separation fares best with non-prototypical, unproductive patterns, which are thus excluded from the core mechanism of morphology.

A derivational paradigm is the systematic synchronic core of a **word family** (in the traditional sense, not in the definition of Corbin (1997b), see below under WF nest). Thus a derivational microclass can be formed from derivational paradigms, whereas no meaningful classes can be formed from word families. Hierarchically inbetween the concepts of word family and WF paradigm, we can situate the concept of **WF nest** (G Wortbildungsnest < R slovoobrazovateľnoe gnezdo), which consists of all synchronic and morphosemantically identifiable derivations from the same lexical base.<sup>20</sup> In contrast, a (traditional) word family contains also derived words which are not connected synchronically (but only diachronically) and have no identifiable morphosemantic relation which conforms with the WF meaning of the respective WFR.

Sometimes both WF paradigms and the microclasses being constituted by them consist just of one WFR, i.e., the sets (of sets) of a basis with one derivative, as is the case with Hungarian unproductive deverbal medium reflexives in *-u/ül*:

(80) gur-ul 'rolls', bor-ul 'turns over', perd-ül 'spins'

The bases are bound roots which reoccur in causative formation such as:

(81) gur-ít, bor-ít, perd-ít 'makes roll/turn over/spin'

The distribution of productivity within a derivational microclass is not only important for the productivity criterion of shift (section 5 (d)) but also for the relation between actual and potential words. For, if there exist two competing productive WFRs, then the ratio of actual to potential words derived by them must decrease, at least tendentially, for both of them, because they constrain each other's actual application (WF synonymy).

Further up in the class hierarchy are: a **derivational class** which consists of those microclasses which have the same base category and output category, the same WF meaning or the same basic/hierarchical superordinate WF

<sup>20</sup> Cf. Uluxanov (1977, 63); Barz (1988, 96). Corbin's (1997b, 9) definition of a "morphological family" is very similar.

meaning.<sup>21</sup> Thus the German derivational class of denominal adjectives (cf. Pounder 1987) consists of the microclass (with relational meaning) of the above-mentioned WF paradigms of *Kind* → adj. *kind-lich* and *kind-isch*, etc., and of the microclass of WF paradigms,<sup>22</sup> such as:

- (82) Stein 'stone' → stein-ig 'ston-y' and stein-ern 'made of stone', Glas 'glass' → glas-ig, gläs-ern, etc.

Another example is the microclass of Russian attenuative adjectives (Zemskaja 1973, 203, 205, 281), as in:

- |      |               |       |             |        |              |        |
|------|---------------|-------|-------------|--------|--------------|--------|
| (83) | krasn-yj      | 'red' | sin-ij      | 'blue' | dobr-yj      | 'good' |
|      | krasn-en'k-ij |       | sin-en'k-ij |        | dobr-en'k-ij |        |
|      | krasn-ovat-yj |       | sin-evat-yj |        | dobr-ovat-yj |        |

This microclass is then extended into a derivational class by the addition of intensifications, as in:

- |      |              |             |              |
|------|--------------|-------------|--------------|
| (84) | krasn-ušč-ij | sin-jušč-ij | dobr-jušč-ij |
|      | krasn-enn-yj | sin-enn-yj  | dobr-enn-yj  |

For the derivational class of nominal alteratives in Italian (consisting of diminutive, augmentative and pejorative formation) see Dressler–Merlini Barbaresi (1994).

If a derivational class comprises two antagonistic productive WFRs, such as Italian diminutives and augmentatives, then antonymic analogy may increase the number of actual derivations (WF antonymy).

The next-higher concept in the hierarchy which represents a valid generalization over WF potentials of word subclasses is a **formal set of microclasses**: this concept comprises all microclasses which have the same sets of lexical bases and the same output categories, e.g., the set of all Hungarian verbs that (at least potentially) allow the same frequentative suffixes, the same reflexive suffixes, the same causative suffixes, and the same possibilitive suffix.

<sup>21</sup> Called *slovoobrazovatel'naja kategorija* 'WF category' in the Russian tradition (Zemskaja 1992, 25–7).

<sup>22</sup> For complications in case of overlapping microclasses see our pre-publication, Dressler–Ladányi (1998, 33).

A **formal set of classes** comprises then all microclasses formed from all bases (including different sets of bases) of the same word class,<sup>23</sup> such as Hungarian frequentatives, reflexives, causatives, possibilitives. The distribution of productivity within the “vertical” and “horizontal” dimensions of this hierarchy is the basis for dealing with morphological richness and complexity, as discussed in the next session.

## 9. Typology of morphological richness and complexity

It is generally acknowledged that inflectional morphology plays a greater role in agglutinating languages than in inflecting-fusional languages and is therefore also considerably richer (cf. Pöchtrager et al. 1998). We postulate the same for derivational morphology (cf. Ülkü 1980 for Turkish). First, we define morphological richness in terms of productive morphological patterns, i.e., in terms of productive morphological categories, rules and paradigms. For example, the following categories are expressed by productive WFRs in Hungarian but are not common in the average European inflecting-fusional languages: causative, possibilitive, active and passive reflexive. The real WF richness of Hungarian comes out only if we look at rule competition and WF paradigms, i.e., at the hierarchical paradigmatic organization of WF (cf. section 8).

Great vertical depth of such derivational class hierarchies and a considerable number of competing WFRs on each class level seems to be typical for agglutinating languages, in contradistinction to many inflecting-fusional languages, such as the Germanic and Romance languages, even if one takes prefixation into account. Slavic languages, such as Russian, however, appear to approach agglutinating languages in the richness, hierarchical depth and rule competition of derivational morphology.

Thus agglutinating languages are richer in both inflectional and derivational morphology than most inflecting-fusional languages.

From morphological richness we pass over to its hyperonym morphological complexity. Morphological complexity, we propose, contains all the morphological patterns of a language, both productive and unproductive ones. Unproductive patterns are lexically stored (according to realistic models of the

<sup>23</sup> Sometimes even subsets of a word class are relevant, such as human nouns or animate nouns instead of the whole word class of nouns.

mental lexicon) and thus do not belong to the active mechanism of morphology which constitutes morphological richness. Morphological complexity must be calculated in terms of the total learning effort devoted to morphology, sc. it contains also the unproductive patterns of a language.

The ideal type of an inflecting-fusional language type has an inflection which is both rich and very complex. The ideal type of an agglutinating language, however, has a very rich morphology, which corresponds perfectly to the much larger role that morphology plays in respect to syntax in this language type, when compared with inflecting-fusional languages. But there are, ideally, no unproductive morphological patterns in agglutinating inflection which would extend complexity beyond richness. However, in WF, also agglutinating languages, such as Hungarian, have many unproductive rules. This fits to the universally greater role of productivity in inflection than in WF.

As predicted, Hungarian which is less typically agglutinating by having some ingredients of inflecting-fusional morphology, has more unproductive patterns both in inflection and in WF than, for example, Turkish (cf. Ülkü 1980).

## 10. External evidence from poetic licence

In previous studies on poetic license Dressler (1981; 1993; cf. Christofidou 1994) has shown that it is more audacious for a writer to produce a poetic occasionalism with an unproductive WFR than with a productive WFR. Here we want to extend this argumentation to our gradation of WF productivity. Our prediction is that the less productive a WFR is, the more audacious a poetic occasionalism (nonce word) formed via this WFR should be. This we will illustrate first with German examples, then with Hungarian ones.

**10.1.** Our first example is German feminine vs. masculine motion. Whereas the formation of feminine counterparts of masculine animal names via the suffix *-in* is productive (see section 4), the inverse formation of masculine counterparts of feminine animal names via the suffix *-erich*, as in fem. *Gans* 'goose' → masc. *Gäns-erich* is of little productivity (if at all). As predicted we found several poetic occasionalisms with feminine motion (see also section 4), e.g.,

(85) Stier → Stier-in 'bull-ess' (Arno Schmidt)

but we found just two examples of masculine motion:

- (86) Krähe 'crow' → Kräh-erich (Joachim Ringelnatz), Wanze 'bug' → Wanz-erich (Heinrich Heine)

**10.2.** Our second example comes from German denominal agent nouns in b-productive *-er* vs. c-productive *-ler*. As expected, there exist many poetic occasionalisms with the first suffix, but the only one with the second in our sample, is:

- (87) Rückstand 'backwardness' → ihr Rückständ-ler! 'you backward people!' (Peter Handke: *Publikumsbeschimpfung*)<sup>24</sup>

**10.3.** Our third example is the derivational class of German denominal adjectives. As we have shown, the suffixations in *-ig*, *-isch*, *-lich* apply to foreign words and are thus type-b (or type-a) productive. In contrast, the suffix *-en/-ern* neither applies to foreign words nor to abbreviations, but just to normal native bases (type-e productivity). As predicted, we find many poetic occasionalisms with the type-b productive suffixes in many authors of our sample (e.g., the Austrians Johann Nestroy, Herrmann Broch, Heimito von Doderer, and the Swiss Max Frisch), but very few with the type-e productive *-e(r)n*, e.g.,

- (88) zwetschk-ern 'made of plums' (Rudolf Bayer), bagatell-en 'like a trifle' (Herbert Rosendorfer), haselnuss-en 'made of hazelnut wood' (Peter Rosei)

However, there are very many such neologisms in the works of the very audacious German author Arno Schmidt, e.g.,

- (89) Er stand... präsidant-en, monument-en, potentat-en, iguanodont-en 'He stood... like a president, like a monument, like a potentate, like an iguanodont' (Arno Schmidt: *Aus dem Leben eines Fauns*)

Other examples of derivations from native and non-native bases are:

- (90) fahrplan-en, kursbuch-en, wasserpflanz-en; bonbon-en, skelett-en, aquarell-en, bajader-en, paysag-en 'like a schedule, course book, water plant; like a bonbon, skeleton, water-colour painting, bayadere, landscape (= French guestword *paysage*)'

<sup>24</sup> But *Rückstand* usually means 'residue' and Handke rather forms it analogically from the adjective *rückständig* 'reactionary', thus semantically not via productive rule formation. Also an analogy to *Mittelständ-ler* 'middle class man' is conceivable.

where even a non-integrated guestword is taken as a base, which flouts a principle of type-a productivity.

**10.4.** Our first Hungarian example is denominal adjective derivation in *-s* and *-i*. Both of them are b- (or a-) productive, but *-i* is less productive because it cannot be attached to abbreviations (c-productivity). For poetic occasionalisms with the first suffix we found many examples in many authors. For the second suffixation we found fewer and most of them in the audacious poet Péter Dobai, e.g.,

- (91) kódex ‘codex’ → kódex-es vs. kozmosz ‘space’, gleccser ‘glacier’ (G Gletscher), inga ‘pendulum’ → kozmosz-i, gleccser-i, inga-i.

**10.5.** Our second Hungarian example is the contrast between two deverbal causative suffixations: *-ta/et* is b-productive, whereas *-a/et* is only e-productive. As predicted, we found many occasionalisms with the first suffix, but very few with the second, e.g.,

- (92) meg-hőköl-tet ‘PFV + shrink back + CAUSE = make somebody shrink back’ (Kornél Döbrentei) vs. rejt-et ‘make somebody hide’ (István Bella).

**10.6.** Our third and last example is passive suffixation with *-ta/etik* (b-productive) vs. *-a/etik* (only e-productive). Again, we found many occasionalisms with the first suffix, but very few with the second, e.g.,

- (93) húroz-tatik ‘be stringed’ (Kornél Döbrentei) vs. el-vár-atik ‘PFV + expect + PASS = be expected’ (György Petri)

With these examples from poetic license we hope to have offered a representative sample of one type of independent external evidence for our gradation of productivity.

## Conclusion

In conclusion, we hope to have shown that a purely grammatical conception and gradation of WF productivity is feasible and fertile, both for theoretical, including typological, purposes and for descriptive purposes. In this sense,

then, morphology by itself has productive grammatical morphology as its core. This holds both for the syntagmatic and the paradigmatic axis of morphology.

Beyond this primary aim of our paper, we want to insist on the basic identity of grammatical productivity in inflection and word formation, which is further evidence both against split morphology of any sort and against Beard's (1995) separation thesis, because productivity holds for both meaning and form of a WFR. The separation thesis would allow the two following conceivable constellations which a semiotically based, non-separationist model disallows:

- (a) a language may have morphosemantically productive derivational rules (e.g., of agent formation), but no productive morphotactic rule of suffixation expressing these derived concepts (e.g., of agent formation);
- (b) a language may have no morphosemantically productive derivational rule, e.g., of agent formation, but may have a productive morphotactic rule of suffixation expressing the derived concept, e.g., of agent formation.

In fact, however, degrees of productivity appear to refer to both morphosemantics and morphotactics.

### References

- Alloni-Feinberg, Y. 1977. Linguistic and socio-demographic factors influencing the acceptance of Hebrew neologisms. PhD thesis, University of Jerusalem.
- Anderson, S.R. 1992. A-morphous morphology. Cambridge University Press, Cambridge.
- André, J. 1971. Emprunts et suffixes nominaux en latin. Droz, Genève.
- Aronoff, M. 1976. Word formation in generative grammar. MIT Press, Cambridge MA.
- Baayen, R.H. 1989. A corpus-based approach to morphological productivity. PhD thesis, Vrije Universiteit te Amsterdam.
- Baayen, R.H. 1992. Quantitative aspects of morphological productivity. In: *Yearbook of Morphology 1991*, 109–49. Kluwer, Dordrecht.
- Baayen, R.H. 1994. Productivity in language production. In: *Language and Cognitive Processes 9*: 447–69.
- Baayen, R.H. – Lieber, R. 1991. Productivity and English derivation: a corpus-based study. In: *Linguistics 29*: 801–43.
- Baayen, R.H. – Neijts, A. 1997. Productivity in context: a case study in a Dutch suffix. In: *Linguistics 35*: 565–87.
- Baayen, R.H. – Renouf, A. 1996. Chronicling the Times: productive lexical innovations in an English newspaper. In: *Language 72*: 69–96.
- Barz, I. 1988. Nomination durch Wortbildung. Enzyklopädie, Leipzig.



- Barz, I. 1998. Neologie und Wortbildung: Zum Neuheitswert von Wortbildungen. In: Teubert, W. (ed.): *Neologie und Korpus*, 11–30. Narr, Tübingen.
- Bauer, L. 1983. *English word-formation*. Cambridge University Press, Cambridge.
- Bauer, L. 1997. Derivational paradigms. In: *Yearbook of Morphology 1996*, 243–56. Kluwer, Dordrecht.
- Beard, R. 1995. *Lexeme-morpheme base morphology*. State University of New York Press, Albany.
- Berman, R. A. 1987. Productivity in the lexicon: new-word formation in the lexicon. In: *Folia Linguistica* 21: 425–61.
- Berrendonner, A. – Clavier, V. 1997. Examen d'une série morphologique dite «improductive» en français: les noms dénominaux en *-age*. In: Corbin et al., 35–44.
- Booij, G. 1977. *Dutch morphology*. Foris, Dordrecht.
- Chomsky, N. 1986. *Knowledge of language: Its nature, origin and use*. Praeger, New York.
- Cristofidou, A. 1994. Okkasionalismen in poetischen Texten. Narr, Tübingen.
- Clahsen, H. et al. 1996. Morphological structure and the processing of inflected words. In: *Essex research reports in linguistics* 13: 1–44.
- Corbin, D. 1987. *Morphologie dérivationnelle et structuration du lexique*. Niemeyer, Tübingen.
- Corbin, D. (ed.) 1991. *La formation des mots: structures et interprétations*. Presses universitaires de Lille.
- Corbin, D. 1997a. Entre les mots possibles et les mots existants: les unités lexicales à faible probabilité d'actualisation. In: Corbin et al., 79–90.
- Corbin, D. 1997b. La représentation d'une «famille» de mots dans le Dictionnaire dérivationnel du français et ses corrélats théoriques, méthodologiques et descriptifs. In: *Recherches Linguistiques de Vincennes* 26: 5–37.
- Corbin, D. et al. (eds). 1997. *Mots possibles et mots existants. Silexicales 1*. Université de Lille.
- Coseriu, E. 1975. System, Norm und Rede. In: Coseriu, E.: *Sprachtheorie und allgemeine Sprachwissenschaft*, 11–101. Fink, München.
- Dal, G. 1997. Tous les mots existants sont-ils possibles? In: Corbin et al., 91–100.
- Doleschal, U. 1992. *Movierung im Deutschen*. Lincom Europa, München
- Dressler, W.U. 1981. General principles of poetic licence in word formation. In: Albrecht, J. (ed.): *Logos Semantics, Festschrift Coseriu*, 423–31. Mouton de Gruyter, Berlin.
- Dressler, W.U. 1989. Prototypical differences between inflection and derivation. In: *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 42: 3–10.
- Dressler, W.U. 1993. Word-formation: poetic license. In: Asher, R. – Simpson, J. (eds): *The Encyclopedia of Language and Linguistics*, 5028–9. Pergamon Press, Oxford.
- Dressler, W.U. 1997. On productivity and potentiality in inflectional morphology. CLASNET Working Paper 7. Université de Montréal.
- Dressler, W.U. – Dziubalska-Kołaczyk, K. 1997. Polish inflection classes within Natural Morphology. In: *Bulletin de la Société Polonaise de Linguistique* 53: 95–119.

- Dressler, W.U. – Ladányi, M. 1998. On grammatical productivity of word formation rules (WFRs). In: Wiener linguistische Gazette 62–63: 29–55.
- Dressler, W.U. – Mayerthaler, W. – Panagl, O. – Wurzel, W.U. 1987. Leitmotifs in Natural Morphology. John Benjamins, Amsterdam.
- Dressler, W.U. – Merlini Barbaresi, L. 1994. Morphopragmatics. Mouton de Gruyter, Berlin.
- Dressler, W.U. – Thornton, A.M. 1996. Italian nominal inflection. In: Wiener linguistische Gazette 55–57: 1–24.
- Erben, J. 1993. Einführung in die deutsche Wortbildungslehre. Schmidt, Berlin.
- Erben, J. 1996. Vorstöße oder Verstöße: Versuch einer Einschätzung von A. Kerrs Neologismen. In: König, W. – Ortner, L. (eds): Sprachgeschichtliche Untersuchungen zum älteren und neueren Deutsch, Festschrift H. Wellmann, 1–11. Winter, Heidelberg.
- Erben, J. 1997. Zum Problem der Nominationsvarianten. Bemerkungen zu Benennungsversuchen in Theaterkritiken Alfred Kerrs. In: Barz, I. – Schröder, M. (eds): Nominationsforschung im Deutschen, Festschrift W. Fleischer, 399–406. Lang, Frankfurt.
- Frauenfelder, U.H. – Schreuder, R. 1992. Constraining psycholinguistic models of morphological processing and representation: the role of productivity. In: Yearbook of Morphology 1991, 165–83. Kluwer, Dordrecht.
- Fuhrhop, N. 1998. Grenzfälle morphologischer Einheiten. Stauffenburg, Tübingen (in print).
- Hacken, P. ten 1994. Defining morphology. Olms, Hildesheim.
- Kastovsky, D. 1982. Wortbildung und Semantik. Francke, Bern.
- Kiefer, F. 1998. How to combine functional and formal approaches. Unpublished paper given in Prague (March 1998).
- Kilani-Schoch, M. 1988. Introduction à la morphologie naturelle. Lang, Bern.
- Kirscher-Durand, Ch. 1997. Mots possibles et mots existants dans une langue morte comme le latin. In: Corbin et al., 173–80.
- Köpcke, K.-M. 1993. Schemata bei der Pluralbildung im Deutschen: Versuch einer kognitiven Morphologie. Narr, Tübingen.
- Ladányi, M. 2000. Productivity as a sign of category change: the case of Hungarian verbal prefixes. In: Dressler, W.U. – Pfeiffer, O.E. – Pöchtrager, M.A. – Rennison, J.R. (eds): Morphological analysis in comparison (Current issues in linguistic theory, 201), 113–41. John Benjamins, Amsterdam.
- Lima, S.D. – Corrigan, R.L. – Iverson, G.K. (eds) 1994. The reality of linguistic rules. John Benjamins, Amsterdam.
- Marle, J. van 1985. On the paradigmatic dimension of morphological creativity. Foris, Dordrecht.
- Marle, J. van 1988. On the role of semantics in productivity change. In: Yearbook of Morphology 1988, 139–54. Kluwer, Dordrecht.
- Marle, J. van 1992. The relationship between morphological productivity and frequency: a comment on Baayen's performance-oriented conception of morphological productivity. In: Yearbook of Morphology 1991, 151–63. Kluwer, Dordrecht.
- Matussek, M. 1994. Wortneubildung im Text. Buske, Hamburg.

- Mayerthaler, W. 1977. Studien zur französischen und theoretischen Morphologie. Niemeyer, Tübingen.
- Moravcsik, E.A. 1975. Verb borrowing. In: Wiener linguistische Gazette 8: 3–31.
- Moravcsik, E.A. 1978. Language contact. In: Greenberg, J.H. (ed.): Universals of human language, 95–122. Stanford University Press, Stanford.
- Motsch, Wolfgang. 1981. Der kreative Aspekt in der Wortbildung. In: Lipka, L. (ed.): Wortbildung, 94–118. Wissenschaftliche Buchgesellschaft, Darmstadt.
- Osman, N. 1994. Kleines Lexikon untergegangener Wörter. Beck, München.
- Petralli, A. 1996. Neologismi e nuovi media. CLUEB, Bologna.
- Pilhak, L. 1975. Das Adjektiv in den Dichtungen Georg Trakls: Untersuchungen zur Syntax, Wortbildung und Semantik. PhD thesis, University of Innsbruck.
- Pinker, S. – Prince, A. 1994. Regular and irregular morphology and the psychological status of rules of grammar. In: Lima, S.D. et al. (eds), 321–51.
- Plag, I. – Dalton-Puffer, C. – Baayen, H. 2000. Morphological productivity across speech and writing. Ms.
- Pöchtrager, M. – Bodó, Cs. – Dressler, W.U. – Schweiger, T. 1998. On some inflectional properties of the agglutinating type, illustrated from Finnish, Hungarian and Turkish inflection. In: Wiener linguistische Gazette 62–63: 57–92.
- Poitou, J. 1997. Régularité, analyses quantitatives et productivité: quelques remarques. In: Corbin et al., 221–30.
- Pounder, A. 1987. Systemangemessenheit in der Wortbildung am Beispiel desubstantivischer Adjektivableitung im Deutschen. PhD thesis, University of Vienna.
- Rainer, F. 1988. Towards a theory of blocking: The case of Italian and German quality nouns. In: Yearbook of Morphology 1988, 155–85. Kluwer, Dordrecht.
- Rainer, F. 1993. Spanische Wortbildungslehre. Niemeyer, Tübingen.
- Rainer, F. 1996. Copernicano e Luterano: sul ruolo del latino nella derivazione deantropo-nimica italiana. In: Lingua Nostra 57: 49.
- Rainer, F. 1997a. Vers une contrainte sur les stades dérivationnelles virtuelles. In: Corbin et al., 231–40.
- Rainer, F. 1997b. Allomorfia e analogia: le condizioni d'impiego dell'allomorfo in *-u-* (tipo *man-u-ale*) in latino e in italiano. Ms.
- Sánchez Miret, F. – Koliadis, A. – Dressler, W.U. 1997. Connectionism vs. rules in diachronic morphology. In: Folia Linguistica Historica 18: 149–82.
- Sandra, D. 1994. The morphology of the mental lexicon: internal word structure viewed from a psycholinguistic perspective. In: Language and Cognitive Processes 9: 227–69.
- Taylor, J.R. 1991. Linguistic categorizations: prototypes in linguistic theory. Clarendon Press, Oxford.
- Temple, M. 1996. Pour une sémantique des mots construits. Septentrion, Villeneuve d'Ascq.
- Thymé, A. – Ackerman, F. – Elman, J. 1994. Finnish nominal inflection: paradigmatic patterns and token analogy. In: Lima, S.D. et al. (eds), 445–66.

- Ülkü, V. 1980. Affixale Wortbildung im Deutschen und Türkischen. Ankara Üniversitesi Basımevi.
- Uluxanov, I.S. 1977. Slovoobrazovatel'naja semantika v russkom jazyke. Nauka, Moskva.
- Wellmann, H. 1975. Deutsche Wortbildung: Typen und Tendenzen in der Gegenwartssprache II: Das Substantiv. Schwann, Düsseldorf.
- Wurzel, W.U. 1984. Flexionsmorphologie und Natürlichkeit. Akademie-Verlag, Berlin. [English translation: 1989. Inflectional morphology and naturalness. Kluwer, Dordrecht.]
- Wüster, E. 1979. Einführung in die allgemeine Terminologielehre und Terminologische Lexikographie. Springer, Wien.
- Zemskaja, E.A. 1973. Sovremennyj russkij jazyk. Proveščenie, Moskva.
- Zemskaja, E.A. 1992. Slovoobrazovanije kak dejatel'nost'. Nauka, Moskva.
- Zemskaja, E.A. (ed.) 1996. Russkij jazyk konca XX stoletija (1985–1995). Jazyki Russkoj Kul'tury, Moskva.
- Zemskaja, E.A. – Kitajgorodskaja, M.V. – Širjaev, E.N. 1981. Russkaja razgovornaja reč. Nauka, Moskva.

Address of the authors: Wolfgang U. Dressler  
Institut für Sprachwissenschaft  
Universität Wien  
Berggasse 11  
1090 Wien  
Austria  
wolfgang.dressler@univie.ac.at

Mária Ladányi  
Általános és Alkalmazott Nyelvészeti Tanszék  
Eötvös Loránd Tudományegyetem  
P.O.Box 107  
1364 Budapest  
Hungary  
ladanyi@isis.elte.hu

## SEMANTIC DIFFERENCES OF SUFFIXAL ALTERNATES IN HUNGARIAN

LÁSZLÓ ELEKFI

### Abstract

Some of the inflectional paradigms of the Hungarian lexicon contain suffixal alternates which do not vary freely with each other, but certain inflected forms are connected to specific senses of the lexical item. Among verbal suffixes the *-ik* suffix marking the third person singular form of the present tense indicative behaves in this way, which on certain verbs can simultaneously signal intransitivity, while the alternate for with the zero suffix is transitive. On other verbs, the choice between the bare *-t* and *-vtt* forms of the past tense suffix (in 3rd person singular) can indicate sense differences. The inflectional paradigm of some verbs can even contain more than one instance of semantic distinction among alternates. A peculiar difference between the multiple plural forms of nouns derived from adjectives is based on whether the nominal meaning is an occasional one or the nominal lexeme has become completely independent. The same difference is reflected in the accusative form of these nouns, too. Different personal possessive suffixal forms of nouns can signal differences between the types of the possessive relation (real possession or part-whole relation), instead of differentiating senses of the basic form. The adverbial forms of adjectives with *-lag*, *-an* and *-ul* suffixes can also indicate sense distinctions.

1. In 1994 the *Dictionary of Hungarian Inflections* [*Magyar ragozási szótár*] was published by the Linguistic Institute of the Hungarian Academy of Sciences in Budapest. The introductory part of this dictionary contains English explanations together with an overview of 6 basic verbal and 6 basic nominal inflectional paradigms of Hungarian. The whole paradigm of the verb contains 61 grammatical forms (infinitives, participles, gerunds and a verbal noun as a derivative in addition to the verbal forms in a strict sense). The nominal paradigm contains 33 forms (nominal cases and other suffixed forms with grammatical roles), which more or less depend on the root of the word. (Several other forms are also derivable, but with such a regularity that there is no need for distinguishing between them according to root forms.) Among the six verbal patterns the dictionary entry corresponds to the third person singular

declarative form of the verb (verbs without the suffix *-ik*), while in the other three the root is followed by a personal suffix *-ik* (verbs ending in the suffix *-ik*). The three patterns represent the three classes of vowels, namely velar, palatal and labiopalatal. Many variants of suffixes harmonize with the types of vowels in the root. Among the six basic forms of nominals three belong to nouns ending in vowels and three ending in consonants. The three classes in each group also differentiate according to classes of vowels.

The 61 verbal and 33 nominal forms of actual words, however, do not always follow the basic patterns, but differ from them in certain respects. There exist some typically alternating roots as well. Taking the characteristic types of alternations into account, the *Dictionary of Hungarian Inflections* distinguishes between 153 verbal and 404 nominal inflectional paradigms. It classifies these paradigms according to the most characteristic suffixal and root-alternations. According to this, the verbal patterns constitute ten groups without *-ik* and nine with *-ik*, while the nominal patterns fall into ten groups of nominal paradigms, ten groups of adjectival paradigms, and a further ten groups the other nominal patterns (numerals and pronouns) and nouns with an incomplete inflectional paradigm, and six groups of adverbials and sentence-words (usually with very limited possibilities for inflection).

In the dictionary, the code of each paradigm consists of two or three elements: it starts with a number, which stands for one of the 19 verbal and 36 other (mainly nominal) paradigm types, followed by a letter standing for the vowel class (*a*, *b* and *c* for verbs and *A*, *B*, *C* and *D* for the other classes of words, the latter comprising the class of words with unstable vowel harmony), and if the paradigm is not a main type within the group, the letter is followed by another number.

The morphological system of Hungarian is so rich (also in forms with grammatical function, and their freer and more bound varieties as well) that the *Dictionary of Hungarian Inflections* is only a comprised summary of the detailed lexicographic database which contains its entries. (This dictionary was originally intended as a supplement to the *Concise Hungarian Explanatory Dictionary* [*Magyar értelmező kéziszótár*].) The more comprehensive database, entitled *The Morphological System of Hungarian Vocabulary* [*Szókincsünk nyelvtani alakrendszere*] can be found at the Department of Lexicography of the Research Institute for Linguistics of the Hungarian Academy of Sciences in four file-drawers and 240 typed pages of code tables. The entries contained here are the same as the entries in the *Dictionary of Hungarian Inflections*, as well as its system of classifying forms, but the number of sub-

types within the groups is much larger, since it contains 1764 paradigms. The recording of this database in a computerized dictionary form is in progress.

2. In Hungarian it is possible to find polysemous lexemes which show a distinction between their senses only in certain paradigmatic forms. (Forms with different argument structures are also considered as belonging to different senses of the lexeme.) A minimal alternation can be observed in verbal paradigm (1a4), in which the past participle can take both *-t* and *-tt* derivational suffixes in its purely participial meaning (denoting the occurrence or perhaps the result of a particular activity), but if it does not have an argument characteristic of the verb, i.e., it is more of an adjectival type, its suffix can only be *-tt* preceded by a linking vowel, e.g., *géppel írt* (or *írott*) *levél* 'typed letter', but only *írott alkotmány* 'written constitution', *írott jog* 'written law', since here the adjective is more likely to denote a property or a quality. (In the *Dictionary of Hungarian Inflections* only four basic verbs, 29 prefixed verbs and one compound verb belong to this category.)

The property of certain classes of words that they can take the inflectional suffix *-ik* in the third person singular form when they are used unaccusatively, but lack the suffix *-ik* when they are used accusatively, is related to the issue of argument structure. This property is characteristic of the verbs *meggyón(ik)* 'confess.3sg' and *körülfol(ik)* 'flow round'. Their other paradigmatic forms follow the paradigm without *-ik*. A lot of verbs formed with the derivational suffix *-z* have this feature, too, e.g., (4a1) *bifláz(ik)* 'cram', (4b1) *pitiz(ik)* 'beg', (4c1) *lövöldöz(ik)* 'shoot aimlessly', (5a3) *szaglász(ik)* 'nose around', and (5b3) *csempész(ik)* 'muggle'. Some of the ones used more often with the suffix *-ik* have variants without this suffix, which are only used transitively. The variant *gyón* 'confess' of the (11a) verb *gyónik* 'confess', can only occur with an object. A variant of the (11a2) form *múlik* 'pass', in the verb *felülmúl* 'surpass' can only occur with this preverb and an object. Similarly, the rare forms *toj* 'lay (eggs)' of *tojik* 'lay eggs' from category 11a7, *ell* 'bear' of *ellik* 'bear' from category 13b4, and *fos* 'shit' of *fosik* 'shit' from category 14a8 can only appear with an object. – The (11a1) verb *gyűlik* 'gather', can only be used in the form without *-ik* in the presence of an agent subject (referring to a group of living beings or persons) (e.g., *Vérszagra gyűl az éji vad* (Arany) 'The night game gathers at the smell of blood'.) The form of *okádik* 'vomit' without *-ik* also tends to require an object, e.g., *füstöt okád* 'eject smoke'. – The same applies to types 14a1 and 14b1 formed with the derivational suffix *-z*, which only differs from 4a1 and 4b1 in that the variant with *-ik* is the more usual one (that is why the headword also has the suffix *-ik*), e.g., *cuctlizik* 'suck (a rubber

teat)', *verklizik* 'play (a hurdy-gurdy)', and 300 others, which are mostly used without *-ik* when they have an object. – The third person dictionary form *esz* of (19b9) *eszik* 'eat', can only be used with an object. If there is no overt object present in the sentence, the understood object is either first or second person, e.g., *esz a méreg* 'anger is eating (you or me)'. The other forms of the verb, however, only tend to attract the (non-specific) object, e.g., the expression *eszek egy falat kenyeret* 'I eat a bite of bread' can be considered only slightly better than the intransitive form *eszek* 'I am busy eating', which characterises the uneducated speech variety.

The different variants of the factitive suffix can also express sense differences: the verb *él* 'live' (1b4) is the base form of *éltet* (*valakit* 'somebody.acc'<sup>1</sup>), meaning 'let a person live, nurture him/her' and *élet* (*vmely időt* 'a certain time.acc') 'make somebody live a certain kind of time, life'. The latter derivational suffix mainly occurs in prefixed verbs, such as *leültet* (*valakit* 'somebody.acc'), 'make somebody sit down' and *leület* (*bizonyos időt* 'a certain time.acc') 'make somebody stay in jail for a certain time' (1c5). These derivative forms, however, can be considered as independent verbs with a whole paradigm. Similarly, the factitive forms of (2b5) *lép* 'step' mainly express a different meaning in its prefixed forms, e.g., *átléptet* (*valakit* 'somebody.acc') 'make somebody step over somewhere', and *átlépet* (*valamit* 'something.acc') 'make somebody step over something'. Besides the 12c5 form *köp* 'spit' there is *köptet* (*valakit* 'somebody.acc') 'make somebody spit' and *(ki)köpet* (*valamit* 'something.acc') 'make somebody spit (out) something'. The factitive form of the 7a4 *kotor* 'scoop' can be *kotrat* (*valamit* 'something.acc') 'make somebody scoop something' and *kotortat* (*valakit* 'somebody.acc') 'make somebody scoop' as well. The 7a8 *háborog* 'grumble', can either be turned into *háborgat* 'disturb' or *háborogtat* 'make somebody grumble', which have totally different meanings, while *csikorgat* 'grind' (one's teeth) is not completely the same as *csikorogtat* 'make squeak' (an object or a tool), either. – From (11a) *gyónik* 'confess', we get the two different forms *gyóntat* (*valakit* 'somebody.acc') 'make somebody confess' (used with or even without an object) and *gyónat* (*valamit* 'something.acc') 'make somebody confess some sin'.

The sense differences indicated in the examples above are also to be analysed according to whether the different meanings are implicitly (i.e., without any formal sign) present in the basic verb as well, or they are characteristic of the factitive derived forms only, and the different varieties of suffixes have different roles in altering the basic meaning of the word.

<sup>1</sup> The abbreviation *acc* refers to the accusative case suffix.



There are cases where the different forms only occur in the past tense third person singular form, e.g., *ragadt* 'stuck somewhere' (intransitive) and *ragadott* 'got hold of something (a stick or weapon)'. Among the five verbs in this (2a1) type, three verbs show some (but each a different) kind of meaning alternation. The verb *téved* meaning 'lose the way', or 'be wrong' is used intransitively in all its senses, but its past tense form can only be *tévedt* (*valahová* 'somewhere') if it refers to movement, and *tévedett*, if it refers to an attitude about somebody's opinion. Other similar alternations of the 18 palatal (2b1) verbs with the same inflexional paradigm are not connected to meaning differences in the past tense form. – The past tense form of the verb *támad* meaning 'attack' or 'arise' is *támadott* if it means an activity, and *támadt* if it means the beginning of something. The 13 verbs belonging to this group also show alternations in this particular past tense form, but no meaning difference can be observed in them. In this particular type, the same alternation characterises the past participle forms as well.

3. In various groups of verbs with alternating stems we can find sense differentiations in several members of the verbal paradigm. The participles and participle-like derivatives of group 6a are more likely to refer in the full-stem version to the activity itself, and in the vowel-less version to the abstract concept associated with the activity, e.g., *az úton gyalogoló katonák* 'the soldiers walking on the road' (participle), *jó gyalogló* 'good walker' (participle turned into a noun), *alig bírja a hosszú gyalogolást* 'he is hardly able to walk long', but the form *gyaloglás* 'long-distance walking' is used to refer to a kind of sport. Ten basic verbs belong to this latter group, e.g., *bujdokol* 'hide', *hajol* 'bend', etc. – The full stem of the (8b9) verb *jelez* 'sign' is only used in the sense 'provide with a sign', especially in participle-like derivatives, like *jelező* 'the person signing', *jelezendő* 'to be provided with a sign' and *jelezés* 'provision with a sign'. In the case of the (8c4) verb *köröz* 'circle' the vowel-less stem variant is only used in the deverbal noun form to sign a difference in meaning, e.g., *körzés* 'circling' can only refer to the movement of the body, typically as a gymnastic exercise. – In the 6a3 type represented by *káromol* 'curse' the first person form also tends to show some tendency of division, e.g., *káromolok* 'curse.1sg<sup>2</sup> somebody', *búvárolok* 'research.1sg' something (at the moment), vs.

<sup>2</sup> Verbal suffixes are abbreviated in the text as follows:

|                             |                           |
|-----------------------------|---------------------------|
| 1sg: first person singular  | 1pl: first person plural  |
| 2sg: second person singular | 2pl: second person plural |
|                             | 3pl: third person plural  |

*káromlok* ‘curse.1sg’ and *búvárlok* ‘research.1sg’ (several times, habitually). 6b is similar in this respect, cf. *perelek* ‘sue.1sg’ and *bérelek* ‘rent.1sg’ (now) vs. *perlek* ‘sue.1sg’ and *bérlek* ‘rent.1sg’ (in principle or habitually). A more significant meaning difference can be observed in some forms of the (6c4) verb *becsül* ‘value’, cf. *becsülök* ‘I have a high opinion of (somebody or something)’ and *becslek* ‘I estimate (an object)’. This is the reason why *becsülés* ‘esteem’ is different from *becslés* ‘estimation’. – In the inflectional paradigm of the (8b2) verb *képez* ‘form’ there is hardly any difference between the more concrete meaning of *képezünk* ‘form.1pl’ and the more abstract meaning of *képzünk* ‘form.1pl’, the difference is more significant in the case of the participial forms *képező* ‘the one who forms’ and the adjectival *képző* ‘forming’ (which has also turned into a noun in several of its senses, e.g., [*tanító*] *képző* ‘teacher training college’ and *képző* ‘derivational suffix’), although not so great as that between the participial and nominalised forms of the verb *ebédel* ‘have lunch’ (lacking alternation), which are *ebédelő* ‘the person who eats lunch’ and *ebédlő* ‘dining-room’.

4. Among the nominals let us first mention those nouns which have been nominalised from a suffixed adjective, such as the ones similar to the (1A2) *gyulai* ‘the one from Gyula’ and the (1B2) *debreceni* ‘the one from Debrecen’. Their plural is formed with a linking vowel if it is to denote the people living in or originating from the particular town (e.g. *gyulaiak* ‘people from Gyula’ and *debreceniek* ‘people from Debrecen’), but without a linking vowel if it refers to the goods (mainly types of food) named after the town (e.g. *gyulaik* and *debrecenik*, denoting particular types of sausages). The type denoting the people is formed regularly from all placenames, but the names of products are lexicalised, therefore unpredictable.

The reason for the above difference is that nominalisation takes place in stages. One of the first stages following ellipsis is a contractional use, in which the noun ‘person’ qualified by the adjective is missing. A nominal incorporating the meaning ‘person’ is already lexicalised as a noun, but it still bears the mark of occasionality through the linking vowel, which reminds of adjectival declination. This primary stage of nominalisation regularly appears in the paradigm of many words which have dual category membership (adjective and noun). The highest stage of nominalisation is reached when the linking vowel characteristic of adjectival plurals disappears. This kind of nominalisation, however, is not a grammatical but a lexical phenomenon (Tompa 1957).

Among consonant-final nouns the ones formed with the derivational suffix *-s* behave almost similarly. At the first stage of nominalisation their

plural suffix is preceded by a low linking vowel, e.g. *okosak* 'clever thing.pl<sup>3</sup>' (e.g., *okosakat mond* 'he says clever things'). Here the sense denoting the person is more lexicalised, appearing with a plural suffix preceded by a mid linking vowel different from that of the adjective (e.g. *így mondják az okosok* 'this is how clever people say it'). This difference manifests itself in the accusative form in such a way that the form *okosat* is used to mean 'clever thing' and the form *okost* to mean 'clever person'. – The word *órás* 'sy/sg having to do with watches/an hour' has reached a higher stage in the process of nominalisation. Its accusative form is *órásat* if it is to mean 'something lasting for an hour' (occasionally lexicalised) and *órást* if it is to mean 'a person working with watches'. The above division into adjectival and nominal meaning is sometimes so great that the sense of the word closer to the adjective belongs to paradigm type 2A9 and the one closer to the noun belongs to the type 4A (which lacks a linking vowel in front of the accusative suffix). In the former case the plural form is *órásak* (less frequently: *órások*), in the latter, *órások*. The nominal plural of *piros* 'red' is *pirosok* only in case it refers to a traffic sign or the suit of cards (*áthajtott a pirosokon* 'he crossed the streets while the lights were red', *pirosokat hív* 'to call hearts'). Similarly, *párosok* means 'double games', *tilosok* 'prohibitory signs', *világosok* 'whites' (in chess). The majority of the 106 nominals in group 2A9 are adjectives, and their plural form alternating between *-ak/(-ok)* has no distinguishing power by itself, only bears the possibility of occasional nominalisation. Only 6 words in this group are pure nouns, while 6 further words are nouns according to their second category. Only among the latter ones can we find lexemes distinguished the grammatical forms mentioned above.

The compulsory sense differentiation in the plural forms of the nominals described above (which is also marked in writing) can be observed among the palatals only in those containing lip-rounded vowels. The reason for this is that the distinction between low [e] and mid [ɛ̃] (IPA: [ɛ] and [ɛ̃]) only shows up in the spoken language, and only for some speakers even there (although there can exist a source text for the main dictionary which marks it). In type 2B6, where, apart from some one-syllable nouns like *jel* 'sign', adjectives formed with a derivational suffix *-s* are in the majority, only the alternation between the *-t* and *-et* forms of the accusative affix distinguishes the accusative form of the adjective from that of nouns belonging to the 4B type (the plural form of which is formed with the linking vowel *ë* by those who make such a distinction). Among the 90 nominals which belong here, only 5 are pure

<sup>3</sup> The abbreviation *pl* refers to the plural suffix.

nouns, while 39 are adjectives and 46 are adjectives according to their primary category (and there is also a secondary noun which can be suffixed in this way). The 46 words which are primarily adjectives can be characterised by the 4B paradigm in their secondary (nominal) sense (that is, in written form, the suffix *-t* without a linking vowel is used to form the accusative). Only two words show a sense differentiation within the above grammatical category (and the 2B6 paradigm). The root *jel* 'sign' takes a suffix *-t* if it refers to an activity, movement, sound, i.e., something to be perceived in time (*jelt ad* 'gives a sign'), and takes a suffix *-et* if it is to denote a visible and lasting figure (*tesz ide egy jelet* 'put a mark here', but the form *jelt* can also appear in this sense). The accusative form of the adjective *tizedes* 'decimal' is *tizedest* without alternation if it refers to a decimal fraction, e.g., *vegyesen írt közönséges törtet és tizedest* 'he mixed simple and decimal fractions in writing'. (This use is beginning to lead towards occasional nominalisation.) – Compared to the nouns in class B, the ones in the lip-rounded class (C) show sense differentiation in more forms: similarly to class A, the difference in the quality of the linking vowel is obvious not only in front of the plural suffix but also in front of certain personal suffixes. Among the 19 adjectives which belong to 2C4 only 10 are adjectives primarily. Here the difference is significant compared to the completely nominalized type of *elsős* 'the one from the first (class)', (4C). But the paradigm of *vörös* 'red', a secondary noun, shows a differentiation according to which sense of the word is used, e.g., the forms *vöröset* 'red.acc', *vörösök* 'red.pl', *vörösöm* 'red.1sg.', *vörösöd* 'red.2sg', can refer to the soldier of the Red Army, etc., while the forms *vöröset* 'red.acc', *vörösek* 'red.pl', etc. refer to wine. The sense of the adjective *vörös* which refers to the person has advanced more in the process of nominalisation than the sense referring to wine.

Among the nouns classified into type 4A5 (on the basis that they mainly take the accusative suffix without the linking vowel), there are also some showing the same kind of distribution as indicated above. Some of the secondary nouns takes a bare *-t* inflectional suffix if it refers to a person or a ghost, e.g. *gonoszt* 'evil.acc', but an *-at* suffix if it refers to an object (since this use has not totally broken away from the adjective), e.g., *gonoszat cselekedett* 'he did something evil'. The noun *lapos* 'flat' does not refer to a person but to a place, lowland, while the form *laposat* 'flat.acc' in *laposat pislant* 'have lids as heavy as lead' is an adverbial recently nominalized from an adjective. – Several among the adjectives inflected in this way show a difference: the sense denoting a person is closer to becoming a noun, and occurs more often with a bare *-t* suffix, cf. *adományost* 'beneficiary.acc' vs. *adományosat* 'sg about donation.acc', *kommencióost* '(person) serving for payment in kind.acc', vs. *kommencióosat*

(also possible with *-t*) '(sg) according to payment in kind.acc', *robotost* 'bondman.acc' vs. *robotosat* 'sg spent working.acc', and *tanyást* '(person) living at a farm.acc' vs. *tanyásat* '(something) which is associated with a farm.acc'.

The compound forms of the adjective *gyilkos* 'who/that kills' show a difference of meaning in the plural, e.g., *öngyilkosok* is '(people) committing suicide' but *öngyilkosak* means '(things) causing their own destruction', *testvérgyilkosok* is '(people) murdering their brothers or sisters' but *testvérgyilkosak* means '(things) causing the death of a brother or sister'. The adjective *talpas* 'having a foot' also behaves similarly, since *talpasak* denotes types of glass, goblets, while infantry men are referred to as *talpasok*.

The differentiation described above also exists among the members of types which can take adjectival suffixes. The adjective *aktív* 'active' from type 12A6 takes a mid linking vowel if it refers to regular troops (e.g., *behívtak néhány nyugdíjas kollégát és sok aktívot* 'they have called in some retired colleagues and many present employees', *ezek az emberek már nyugdíjasok, nem aktívak* 'these people are already retired, not working'), otherwise, mainly referring to things, the linking vowel is low, e.g., *ezek az erők most aktívak* 'these forces are now active'. Both plural forms of *impulzív* 'impulsive', namely *impulzívak* and *impulzívakok* can be used for people, but only the former can be used to characterise certain effects as well. People can be surly permanently, which is expressed by the form *morcok* 'surly.pl' or temporarily, expressed by *morcak* 'surly.pl', times, however, can only be referred to by the latter. People can be happy permanently, which is expressed by the form *boldogok* 'happy.pl' and temporarily, implied by the form *boldogak* 'happy.pl'.

Among the adjectives which can appear with the bare accusative suffix *-t*, some in type 15A3 show a meaning differentiation between two possible forms, cf. *konyhást* '(person) helping at the kitchen.acc' vs. *konyhásat* '(something) provided with a kitchen.acc', *laikusok* '(people) who are laymen' vs. *laikusak* 'non-trained.pl', *mániákust* 'maniac.acc' vs. *mániákusat* 'fan.acc', *méltóságosat* 'dignified.acc' vs. *méltóságost* 'Honorable.acc', *misztikusok* 'mystical.pl' vs. *misztikusak* 'mysterious.pl', *muzsikusok* '(people) making music' vs. *muzsikusak* '(people) enjoying music', *nádasok* 'reedy.pl' vs. *nádasak* '(things) covered with reeds' and *olaszosak* '(things/people reminding of Italians' vs. *olaszosok* '(people) learning Italian'. The low linking vowel refers more to a characteristic feature, while the mid linking vowel to a classification or qualification. These latter ones are close to becoming a noun. The distinction exists in the following cases as well: (15A8) *nyugdíjast* '(person) getting a pension.acc' vs. *nyugdíjasat* '(sg) involving a pension.acc', *nyilast* '(a person) having an arrow.acc' vs. *nyilasat* '(sg) decorated with an arrow.acc',

*ösztöndíjast* '(person) holding a scholarship.acc' vs. *ösztöndíjasat* '(sg) related to scholarship.acc', (15A12) *lutheránusok* 'Lutheran.pl' vs. *lutheránusak* 'equivocating.pl'. Moreover, compare (15C1) *bűnösök* 'guilty (people)' vs. *bűnösek* 'guilty' (people or things), *közömbösök* 'indifferent (people)' vs. *közömbösek* 'indifferent (situations or materials)', (15C3) *ejtőernyősök* '(people) who are paratroopers' vs. *ejtőernyősek* 'something having to do with a parachute.pl' (e.g. parachute drops).

5. The pronoun *más* 'different, other (thing or person)' in (25A11) can stand for either a noun or an adjective. The forms *mást* 'different.acc' and *mások* 'different.pl' relate primarily to the nominal meaning, but many people do use it in its adjectival sense, e.g., *mást szeret* 'loves somebody different from the person/thing talked about', *mások* 'other people or things', *nem ilyen dobozt kérek, hanem más* (or, rarely: *másat*) 'I want not this kind of box but something of a different quality', *a körülmények most mások* or *mások* 'circumstances are now different'. The forms with a low open linking vowel are always adjectival.

The word *pár* 'pair' from type 4A6 is inflected as *párt* 'pair.acc', *párok* 'pairs', *párom* 'pair.1sgposs'<sup>4</sup>, *párja* 'pair.3sgposs', *párjaim* 'pair.1sgposs.pl', if it refers to an entity which consists of two parts, but if it refers to two pieces chosen temporarily, it is inflected as *párat* 'pair.acc' (e.g. *a kolbászból kérek egy párat* 'I want a pair of these sausages'), *páram* 'pair.1sgposs' (*nekem is van belőle néhány páram* 'I have some pairs of it, too'). Moreover, in this latter use the form *párain* can also occur besides *párjaim* (e.g., *virsliből friss párain vannak* 'I have fresh pairs of sausages'). This form of *pár* with the added *a* is closely related to the numeral *pár* 'several' used as an adjective.

6. A peculiar meaning difference can be expressed by the third person singular form of the personal possessive suffix. In the case of *fonal* 'thread' (type 2A7), the form *fonala* 'thread.3sgposs' expresses close relation, sometimes a metaphorical identification (*az elbeszélés fonala* 'the thread of the story'), while the variant *fonalja* 'thread.3sgposs' can only be used to denote a certain type of commodity owned by somebody. In the *Dictionary of Hungarian Inflections* we forgot to indicate this particular difference, but reference to this feature can be found at many other entries with alternating meanings. The possessive form of the word *talp* 'sole' (type 2A8), can either be *talpa* 'sole.3sgposs', referring to the part or piece of somebody/something, e.g. *az ember, a cipő talpa* 'the

<sup>4</sup> The abbreviations *1sgposs*, *2sgposs*, etc., refer to the personal possessive markers.

sole of a person, of a shoe' or *talpja* 'sole.3sgposs', referring to the shoemaker's findings. The possessive form of the noun *szál* 'thin, long piece' is *szála* 'thin, long piece.3sgposs' when it is used in connection with flowers or hair, while the form *szálja* 'thin, long piece.3sgposs' refers to the pieces of flowers owned by a florist. This form can also be a distributive marker, e.g., *szálja 50 forint* 'a piece (of flower) costs 50 forints'. The addition of the personal possessive suffix to *vonat* 'line' usually results in the form *vonala* 'line.3sgposs', but in the old days the conductor on the tram could be asked for a ticket with the question *Van-e vonalja?* 'Have you got a through ticket for one line?'. (The latter expression was used at a time when there was a conductor selling tickets on the tram and the bus.)

The word exemplifying type 2B5 is *fék* 'brake'. Its variants with *-je* vs. *-e* do not give rise to sense distinction, but the forms of the word *fehér* 'white', one of the six nominals belonging to the same category, show peculiar differentiation. In the sense of 'white colour' the personal possessive suffix *-e(-je)* is used, in the sense of 'white wine' only the *fehére* 'white.3sgposs' form is possible, if it is to mean 'white person', both forms *fehérje* and *fehére* are possible, but in connection with eggs or pigs only the form *fehérje* can occur. (The attachment of the possessive suffix to the stem leads to a different range of problems which will not be touched upon here.) In the case of type 2B8, illustrated by *lék* 'leak', the possessive form of which only rarely contains the *-j* sound, the suffixed forms of the noun *gép* 'machine' show some kind of sense distinction, e.g., in the sentence *Neki is van ilyen gépje* 'He also has such a machine' the less common variety of the suffixed form (mainly found in less educated style, in addition to the common form *gépe* 'machine.3sgposs') is used to express real possession. The buyer's question to the greengrocer when buying a watermelon also sounds as follows: *Ideadja a lékjét?* 'Would you give me the part cut out?'. (This difference was not included in the dictionary of affixation, either.) – The type 2C7 *öl* meaning 'lap' or 'cord', refers in the form *öle* 'lap.3sgposs' to the part of the body (and its metaphorical uses), while it is a measure word in the form *ölje* 'cord.3sgposs'.

In the paradigm group No. 3 (characterised by an accusative suffix with a mid linking vowel) there are also paradigm types allowing the sense division referred to above. The mostly free variation in the third person inflection in paradigm type 3A2 (serving as a pattern for 133 simple nominals) is partly polarized in the inflected forms of the noun *anyag* 'material', e.g., while the expression *valaminek az anyaga* 'the material of something (the material it consists of)' only allows the form without the *-j-*, to refer to the goods a shopkeeper has, both the form *anyaga* 'material.3sgposs' and *anyagja* 'material.3sgposs'

can be used. In the expression *az ég minden csillaga* 'all stars of the sky' the inflected form *csillaga* 'star.3sgposs' can also be substituted for *csillagja* 'star.3sgposs', but only the latter one can be used to refer to the star-shaped insignia that rank soldiers have. For the toys owned by children or shopkeepers both inflected forms of the noun *játék* meaning 'toy' or 'play', namely *játéka* 'toy.3sgposs' or *játékja* 'toy.3sgposs' can be used, but only *játéka* is appropriate to refer to the play of a musician or an actor, that of a child or a sportsman when they play a game, or to the somewhat looser movement of a component part of a machine. The noun *szolgálat* 'service' takes an *-a* personal possessive suffix in all its senses, but the expression *szolgálatjára* 'at your service' has reserved an old plebeian variant. – The type 3A4 differs from the former one in that only before the plural suffix can we perceive an alternation, e.g., *korlát* 'barrier', *korlátja* 'barrier.3sgposs' but *korlátaim* 'barrier.1sgposs.pl' or *korlátjaim* 'barrier.1sgposs.pl'. Among the 46 nouns behaving in the same way the forms in the inflectional paradigm of *citrom* 'lemon' show a division of meaning: to refer to the fruit itself, both *citromaim* 'lemon.1sgposs.pl' and *citromjaim* 'lemon.1sgposs.pl' can be used, but only the former one is appropriate to refer to the lemon tree. – We can find similar ones among the 62 nouns which pattern according to 3B2: *bélyege* 'mark.3sgposs' denotes the marks on something or characteristic of something, while the forms *bélyege* 'stamp.3sgposs' or *bélyegje* 'stamp.3sgposs' refer to the stamps owned by somebody. The form *füzete* 'exercise-book.3sgposs' can denote an exercise-book owned by a person, while the form *füzetje* 'exercise-book.3sgposs' is the one appropriate to refer to the shopkeeper's goods with. The latter form can also act as a classifier, expressing that each piece of a publication costs a certain amount of money. The forms *kerete* 'frame.3sgposs' or *keretje* 'frame.3sgposs' can both be used to refer to the frames possessed by the picture framer, but if it is to denote the frame belonging to a picture, only the form *kerete* 'frame.3sgposs' is possible, just like in cases where the noun refers to a more abstract notion.

The body (an essential feature) of a person, an animal or an object can be denoted by the inflected form *teste* 'body.3sgposs', while the form *testje* 'body.3sgposs' refers to a geometrical solid owned by somebody. The noun *tiszt* 'officer', following the 3B6 pattern, is used in the form *tisztje* 'officer.3sgposs' to refer to a person, but it appears without the *-j-* when used in a figurative sense in the following archaic expressions: *tisztében áll* 'it is his duty' and *lerója tisztét* 'perform his duties'. In its most common sense, the noun *cseléd* 'maid' also takes a *-je* affix. The noun *szelep* 'valve' referring to a component part of a machine is inflected as *szelepe* 'valve.3sgposs', but to denote an object owned by somebody, the form *szelepje* 'valve.3sgposs' is used. – The (3C2)



*küszöb* 'threshold' only appears in the form *küszöbje* if it refers to somebody's possession (otherwise *küszöbe* 'threshold.3sgposs').

The above type of meaning distinction characterising nouns with alternating stems can extend to the varieties of the stem as well. The possessive form of *ajtó* 'door' is *ajtaja* 'door.3sg.poss' if it refers to the door belonging to a house, a flat or a car, but it is *ajtója* if it refers to an object possessed by a carpenter. The form *mezeje* 'field.3sgposs' is used to denote the field owned by a farmer, a community or a village (in the original 'arable land' sense of the word). But if the possessive form is used to refer to a field of force or a small surface, then the variant *mezője* 'field.3sgposs' is used. If the word *fő* 'head' is to denote a constituent, ruling part of a person, an object, a group, or an abstract thing, its possessive form is *feje* 'head.3sgposs', but if it refers to the whole person, e.g., in the case when the number of people belonging to a leader or a group are enumerated, the form *fője* 'head.3sgposs' is the appropriate one. The difference is similar between *munkaereje* 'working capacity.3sgposs' and *munkaerője* 'manpower.3sgposs'. If a woman belonging to somebody is referred to by the form *nője* 'woman.3sgposs', then she is not meant to be his wife, only if the form *neje* 'woman-3sgposs' is used. The possessive form of the words *tüdő* 'lung' and *velő* 'marrow' are *tüdeje* 'lung.3sgposs' and *veleje* 'marrow.3sgposs' (or *velője*) if they refer to parts of humans and animals, but *tüdője* 'lung.3sgposs' and *velője* 'marrow.3sgposs' if they refer to goods sold by the butcher. (But in the sense 'the essence of something' only the form *veleje* 'marrow.3sgposs' is appropriate.) The roof belonging to buildings and cars is normally inflected as *teteje* 'roof.3sgposs', but if we wish to refer to the objects in possession of their owners, only the form *tetője* 'roof.3sgposs' is possible. The noun *vessző* meaning 'twig' can both be inflected as *vesszője* 'twig.3sgposs' and *vesszeje* 'twig.3sgposs', but in the sense of 'comma' the only possible possessive form is *vesszője* 'comma.3sgposs'. The same kind of difference can also be observed between the forms *szőlője* 'grape.3sgposs', which refers to the fruit, and *szőleje* 'vineyard.3sgposs', which refers to the land where the fruit is grown. The form *szülei* 'parent.3sgposs.pl' refers to one's parents, while the form *szülői* 'parent.3sgposs.pl' to the parents belonging to a school or community.

The two third person possessive forms of the nominalised adjective *belső* 'inner' are also distinct: *belseje* 'inner.3sgposs' refers to the inner part of something (related forms are *belsejük* 'inner.3plposs', moreover, *belsejű* (adjective) 'having an inner part of some particular quality'), while *belsője* 'inner.3sgposs' denotes the rubber hose or tube belonging to balls and tyres. The difference is similar between the forms *külseje* 'outer.3sgposs', which denotes the outer

part or shape of something, and *külsője* 'outer.3sgposs', which is the outer cover of balls and tyres.

Idiosyncratic stem variants serve the purposes of distinguishing between the senses of the noun *bíró* 'judge'. The plural form *bírák* 'judge.pl', as opposed to *bírók* 'judge.pl', is interpreted only by a few people to mean the judiciary (perhaps the ancestors of the Jewish kings in the Old Testament), but among the possible possessive forms, *bírám* 'judge.1sgposs' denotes the person who judges me, and *bíróm* 'judge.1sgposs' can refer to a judge whom I know, with whom I am in a closer contact.

A secondary meaning is reflected by the (4A2) form *császársa* 'Caesarean section.3sgposs', while in the sense of 'emperor' the form *császára* 'emperor.3sgposs' is used. In the old sense of the word *monitor* 'battleship' both possessive forms *monitora* 'battleship.3sgposs' and *monitorja* 'battleship.3sgposs' are acceptable, but in the sense of 'screen' only *monitora* 'screen.3sgposs' is possible. The alternations described above exemplify individual sense distinctions, which can be observed at several members of the noun category 4A2. No such differentiation, however, is traceable in types 4B2 and 4C2, which contain high vowels. – The *-ja* personal suffix can attach to the nouns *auktor* 'author', *forradalmár* 'revolutionary', *korrepetitor* 'coach', *traktor* 'tractor', *vásár* 'market', *vektor* 'vector', *viador* 'gladiator' from category 4A4, but it is quite common with 10 words ending in *-tan* 'theory' (like *alaktan* 'morphology') to refer to concrete objects, like *alaktanja* 'morphology.3sgposs' referring to a textbook. – There is hardly any meaning differentiation in the third person singular suffixed forms of *tanár* 'teacher', type 4A14, while the contrast is more striking in the case of *titkár* 'secretary': the form *titkára* 'secretary.3sgposs' is used to refer to the person working for a party or an association, while the form *titkárja* 'secretary.3sgposs' to a person working for the boss. The (4A15) *futár* 'messenger' can only be suffixed as *futárja* 'messenger.3sgposs' if it refers to the chess-piece bishop. A similar difference can exist between the forms *pincére* 'waiter.3sgposs' (of a restaurant) and *pincérje* 'waiter.3sgposs' (of the owner of the restaurant). Besides *métere* 'meter.3sgposs', the form *méterje* 'meter.3sgposs' exists as well, and it is taken to refer to one meter from the length of a measurable material. The (4A16) *sor* 'row' is almost always suffixed as *sora* 'row.3sgposs', but the other form is used in the expression *ez a világ sorja* 'this is the order of things'. (Earlier, e.g., by Arany: *Inni kell, ha rád jön sorja* 'You have to drink when it comes to that'.)

Among the 9 root nouns which belong to category 7A5, *barom* 'beast' is suffixed as *barma* 'beast.3sgposs' to refer to somebody's domestic animal, but as *baromja* 'beast.3sgposs' to refer to an uneducated, stupid person. The form

*csokra* 'bunch.3sgposs' is used to denote a bunch of flowers or one belonging to a person, while *csokorja* 'bunch.3sgposs' means 'one bunch of the type of flowers mentioned previously'. The form *cukra* 'sugar.3sgposs' is the one in general use, since *cukorja* 'sugar.3sgposs' can only refer to diabetes or to one piece of candy. The possessive form of *halom* 'pile' is usually *halma* 'pile.3sgposs', but it is *halomja* 'pile.3sgposs' if it is used figuratively, to mean a large quantity, a pile of something. Similarly, the figurative sense of the word is referred to by the 7A6 forms *szatyort* 'bag.acc', *szatyorja* 'bag.3sgposs' meaning 'old woman', which exist beside the regular forms *szatyrot* 'bag.acc', *szatyra* 'bag.3sgposs', from the 5 roots of category 7A6. From among the 4 roots in category 7B2 the language uses only the varieties of *iker* 'twin' for meaning differentiation: *ikret* 'twin.acc' (which is rarely used) refers to a twin brother or sister, while *ikert* 'twin.acc' refers to a telephone station. In addition to the type 7C6 form *csülke* 'knuckle.3sgposs' there exists the rarer form *csülökje* 'knuckle.3sgposs' to refer only to the product.

There are a few stems whose primary senses can be characterised by a low stem-final vowel, and the secondary senses by a mid stem-final vowel. (This distinction is parallel to the one according to which the form of the possessive personal suffix including *-j-* implies a looser relation than the one without it. Possession is considered secondary to direct connection, this is why the forms *talpa* 'sole.3sgposs' and *bőre* 'skin.3sgposs' forms are used to refer to the sole and skin or leather of a person or a shoe, respectively, and the forms *talpja* and *bőrje* to the material owned by a shoemaker, which, however, do not necessarily belong to him.) The (3A6) *sark* meaning either 'heel', 'corner' or 'pole' (more often used as *sarok*, type 7A1, in its primary sense) is inflected as *sarkat* 'heel.acc', *sarkak* 'heel.pl', *sarkam* 'heel.1sg' if it refers to a part of the body, but as *sarkot* 'corner.acc', *sarkok* 'corner.pl', *sarkom* 'corner.1sgposs' if it is used in its second sense, and only as *sarkot* 'pole.acc' and *sarkok* 'pole.pl' if it is used in its third sense. In its original sense, the (3A12) word *szarkaláb* 'the leg of a magpie' would have been inflected as *szarkalábat* 'leg.acc', *szarkalábak* 'leg.pl' and *szarkalábaim* 'leg.1sgposs.pl', but this sense is so rare that the *Hungarian Concise Explanatory Dictionary* (1972) does not even mention it. According to the above dictionary, the primary sense of the word is the name of a flower, common larkspur, and it has become so detached from the original meaning of the compound form mentioned above that its typical paradigm is *szarkalábot* 'common larkspur.acc', *szarkalábok* 'common larkspur.pl', *szarkalábom* 'common larkspur.1sgposs' and *szarkalábja* 'common larkspur.3sgposs'. Its second and third senses ('scrawly handwriting' and 'wrinkles on the face'), though metaphorical, are pictorially closer to the

original meaning (which does not appear in the above mentioned dictionary), their paradigm usually includes the original *-a-* sound, therefore the personal possessive suffixes are *-a* and *-ai*.

The (5A2) *út* 'road' and its several (5A3) compounds can be used with a personal suffix without *-j-* only in the plural form. The sense distinction is also traceable here: the form *utaim* 'road.1sgposs.pl' is used more often to refer to the speaker's travels, since if it is to mean a concrete object, the roads or carriage-ways of a county, the forms *útjai* 'road.3sgposs.pl' and *kocsiútjai* 'carriage-way.3sgposs.pl' are used normally. To refer to the boss's tours of inspection, or the flights of the head of state, both forms *szemleútjai* 'tour of inspection.3sgposs.pl' and *szemleutai* 'tour of inspection.3sgposs.pl', *repülőútjai* 'flight.3sgposs.pl' and *repülőutai* 'flight.3sgposs.pl' can be used, respectively, but the roads for tourists can only be referred to with the possessive form *turistaútjai* 'tourist road.3sgposs.pl'. – The accusative form of the (5B1) *ér* 'blood-vessel' is *eret* 'blood-vessel.acc', but it can be *ért* as well, if it is to mean 'brook'. The stem of the noun *tér* 'space' or 'square' is left unchanged only if it is to mean 'a small gap', and in some phraseological units, e.g., *tért* (or *teret*) *hagy* 'leave room', *tért hódít* 'gain ground', *tért nyer* 'gain ground', and *tért nyit* 'offer a large scope'.

The possessive form of the noun referring to the pigs owned by the farmer is *disznója* 'pig.3sgposs' or *disznaja* 'pig.3sgposs' and *disznói* 'pig.3sgposs.pl', perhaps *disznai* 'pig-3sgposs.pl', but the same word used for the ink-blot made by a child is inflected as *disznója* 'ink-blot.3sgposs'. The inflected form *disznója* can refer to the ace possessed by a card-player, and to a person with obscene speech and disgusting behaviour, who belongs to a company. – The form *apraja* 'small.3sgposs' can refer to the small pieces in some collection, or to the children belonging to a group of people, but the form *aprója* 'small.3sgposs' (meaning 'his change') is only used to refer to the coins owned by somebody. (The anterior constituents of a compound, which become independent as a result of shortening, usually appear unchanged in inflected forms.)

Very rarely the second person form of the possessive suffix can also express a difference in meaning. (Due to space limitations, this type of formal alternation and occasional differentiation could not be included in the *Dictionary of Hungarian Inflections*.) The word *lángész* 'genius' can refer to extraordinary talent (and then the possessive form is *lángeszetek* 'genius.2plposs') or to the talented person himself (*Nektek is van néhány lángészte* 'You also have some geniuses').

A sense differentiation in the forms of nominalised adjectives bearing a *-nként* adverbial suffix is also rare, but it is observable between the form

*apránként* 'gradually' lexicalised as an adverb and the infrequent form *ap-rónként*, which means 'by coins'.

As far as the two nouns in category 8A2 are concerned, the forms *hamvam* 'ash.1sgposs' and *hamva* 'ash.3sgposs' are so rarely used in their original sense, represented by the more common forms *hamum* 'ash.1sgposs' and *hamuja* 'ash.3sgposs', that almost a total meaning split has occurred: the form *hamva*, and especially its plural forms, *hamvaim* 'ash.1sgposs' and *hamvai* 'ash.3sgposs' refer to the ashes of a cremated dead person. The plural form without a personal possessive suffix, i.e., *hamvak* 'ashes', can only be used in the sense 'mortal remains'. The word *odú*, meaning either 'hollow' or 'den' only shows a split of forms caused by grammatical features: *odúja* 'den.3sgposs' refers to the small dwelling of a human or an animal, while *odva* 'hollow.3sgposs' denotes the hollow part of a tree (perhaps the form *odvam* is also possible in the sense of 'my favourite hollow', beside the form *odúm* 'den.1sgposs').<sup>5</sup> – The plural forms of the word (8A9) *mag* 'seed' without a personal possessive suffix show no major sense differentiation: both the forms *magok* 'seed.pl' or *magvak* 'seed.pl' are used, but the latter has a very special shade of meaning, namely 'the seeds of various plants appropriate for propagation'. In the forms containing personal possessive suffixes, the meaning difference is greater: the form *magvam* 'seed.1sgposs' mainly means an offspring (if uttered by a male), while the form *magom* 'seed.1sgposs' can refer to any kind of seed (even sperm). The form *magva* 'seed.3sgposs' is mainly used in a figurative sense, meaning 'offspring', 'centre' or 'essence'.

It would make a long list to enumerate all the simple words which receive different suffixes in compounds in the presence of different anterior constituents, partly because the anterior constituent determines the meaning of the posterior constituent to a certain extent. Some examples, however, which belong to different subgroups in the collection *The Morphological System of Hungarian Vocabulary* seem to be in place here. The different inflected forms of the word *szó* 'word', e.g., *szavak* 'word.pl' vs. *szók* 'word.pl', *szavam* 'word.1sgposs' vs. *szóm* 'word.1sgposs' and *szavunk* 'word.1plposs' vs. *szónk* 'word.1plposs' do not demarcate, although the stem including the *v* is more likely to be used in the sense of continuous speech, while the unchanged stem is in the sense 'one word'. The words contained in the dictionary are referred to by *szavunk* 'word.1plposs' or *szónk* 'word.1plposs' but if we cannot say a word, then the expression *szavunk sincs* 'we are speechless' is used. In third person only the forms *szava* 'word.3sgposs' and *szavuk* 'word.3plposs'

<sup>5</sup> On the split of grammatical forms and partial vs. total word split, see Elekfi (1996).

can appear. The compounds referring to parts of speech, however, denote individual words, therefore the forms *mondatszók* 'sentence-word.pl', *mondatszóm* 'sentence-word.1sgposs' and *mondatszónk* 'sentence-word.1plposs' are more often used than *mondatszavak* 'sentence-word.pl', while in the third person the form *mondatszója* 'sentence-word.3sgposs' is also possible besides *mondatszava* 'sentence-word.3sgposs'. The base form of a derivative, however, can only be referred to by the form *alapszava* 'base form.3sgposs', besides *alapszó* 'base form'. The words *előszó* 'foreword' and *végyszó* 'last word' or 'cue' also behave differently. The word *előszó* 'foreword' refers to an actual text, therefore the form *szó* has the role of a noun of multitude here, and the whole compound is inflected in an idiosyncratic way. Its plural form can only be *előszók* 'foreword.pl', with a personal possessive suffix it becomes *előszóm* 'foreword.1sgposs' (rather than *előszavam*), *előszód* 'foreword.2sgposs', but besides *előszava* 'foreword.3sgposs' we also find the form with the stem unchanged, namely *előszója*. The noun *végyszó* can refer to one word, or more last words. The word or words the actor has to pay attention to are called *végyszó* 'cue' or *végyszók* 'cue.pl', the possessive form of which is *végyszói* 'cue.3sgposs', but used in the sense to refer to the last words of a dying person the same word is inflected as *végyszavai* 'last word.pl'. In this latter sense of the word the form *végyszavak* 'last word.pl' was used previously. The form *nótaszó* 'song' belongs to a different semantic field, since it does not refer to a linguistic unit but to musical sounds, and in this sense it nearly always keeps its original stem in inflected forms. Only in the third person can we find the alternating forms *nótaszava* 'song.3sgposs' vs. *nótaszója* 'song.3sgposs' and *nótaszavuk* 'song.3plposs' vs. *nótaszójuk* 'song.3plposs'. The same behaviour is characteristic of the compounds *harangszó* 'sound of a bell' and *muzsikaszó* 'sound of music'.

Among the inflected forms of the noun *lé* 'liquid' (and in many of its compounds), the forms of the *levet* 'liquid.acc' and *leve* 'liquid.3sgposs' type are more common if the word is to refer to a natural liquid associated with the material of something, e.g., *levet ereszt* 'it gives off juice', *a citrom leve* 'the juice of the lemon'. The stem remains unchanged (less frequently) in forms referring to the artificial liquid owned by somebody, e.g., *iszik valamilyen lét* 'he drinks some juice' or *a büfének elfogyott a léje* 'the canteen has run out of juice'. The stem of the noun *permetlé* 'disinfectant spray', an artificial product, is therefore left unchanged, e.g., *permetlét* 'disinfectant spray.acc', *permetlék* 'disinfectant spray.pl' and *permetléje* 'disinfectant spray.1sgposs' (and only less frequently do we find the form *permetleve* 'disinfectant spray.acc', etc.). The noun *hallé* 'fish soup' patterns the same way. The forms *káposztalé* 'cabbage-pickling brine', *hurkalé* 'pork stock', *húslé* 'gravy' and *mészkénlé* 'cal-

cium sulphite' (although this latter one is not a type of food), however, occur more often with the stem containing *-v-*, e.g., *káposztalevet* 'cabbage-pickling brine.acc' or *káposztalét* 'cabbage-pickling brine.acc', *káposztaleve* 'cabbage-pickling brine.3sgposs' (rarely: *káposztaléje*), while its simple plural form (if it occurs at all) can only be *káposztalevek* 'cabbage-pickling brine.pl'.

The above difference is also relevant in the inflectional paradigm of the (8C1) noun *tő* 'stock'. The form *tőve* 'stock.3sgposs' denotes an inherent part of something, while the suffix *-je* indicates that reference is made to a stock in the possession of somebody. While the form *tőig* 'up to the stock', is a regularly inflected adverb, the form *tővig* 'totally' has gone a long way in the lexicalisation process, e.g., *tővig leégett* 'it has burnt completely'. In addition, the forms *tőn* 'stock.superess<sup>6</sup>' and *tővön* 'stock.superess', *tövet* 'stock.acc' and (less frequently) *tőt* 'stock.acc' and *tővünk* 'stock.1plposs' (*tőnk* 'stock.1plposs') all exist, without any significant difference in meaning. Similarly, the accusative form of *szótő* 'stem' is *szótövet* 'stem.acc' (*szótőt* 'stem.acc'), the plural is *szótövek* 'stem.pl', on the pattern of *tövek* 'stock.pl', while there is only one variant of the forms with the personal suffix, e.g., *szótövem* 'stem.1sgposs', *szótöved* 'stem.2sgposs', *szótöve* 'stem.3sgposs', *szótövünk* 'stem.1plposs', *szótöveim* 'stem.1sgposs.pl', etc. In the inflectional paradigm of the nouns *igető* 'verb stem' and, moreover, *combtő* 'the upper part of the thigh', the ratio of the forms containing an added *v* is also very high. In compounds patterning like *rózsatő* 'rose tree' (e.g., *búzatő* 'wheat stalk', *fartő* 'aitchbone', *naddálytő* 'comfrey', *szőlőtő* 'vine-stock'), however, the inflected form containing an unchanged stem is almost on an equal status with the one containing *-v-*: e.g., *rózsatövet* 'rose tree.acc' vs. *rózsatőt* 'rose tree.acc', *rózsatövek* 'rose tree.pl' vs. *rózsatők* 'rose tree.pl', *rózsatövem* 'rose tree.1sgposs' vs. *rózsatőm* 'rose tree.1sgposs', etc. – The presence of the *-v-* in the whole paradigm is characteristic of the inflected forms of the (8C4) *kő* 'stone', e.g., *kövön* 'stone.superess', *követ* 'stone.acc', *kövek* 'stone.pl', *kövem* 'stone.1sgposs', *kövünk* 'stone.1plposs', etc. In its compound forms, however, we can witness the appearance of the unchanged stem together with the variant containing *-v-*: e.g., the inflected forms of *borkő* 'tartar', *gálickő* 'vitriol', *kilométerkő* 'kilometre stone', *mérföldkő* 'milestone', *rajtkő* 'starting block' can be either *borkőn* 'tartar.superess' or *borkövön* 'tartar.superess', *borkövet* 'tartar.acc' or *borkőt* 'tartar.acc', *borköve* 'tartar.3sgposs' or *borkője* 'tartar.3sgposs', *borkövünk* 'tartar.1plposs' or *borkőnk* 'tartar.1plposs' and *borkövetek* 'tartar.2plposs' or *borkőtök* 'tartar.2plposs'. In these and some other compounds, the same variants

<sup>6</sup> *Superess* denotes the suppressive case suffix.

also exist as in the paradigm of the noun *tő* 'stock', but without any difference in meaning. Some compounds containing the noun *cső* 'pipe' (e.g., *könyökcső* 'angle pipe', *kukoricacső* 'ear of corn' and *látcső* 'telescope') also show the same alternation, just like the word *cső* itself, but the less frequent forms *csőn* 'pipe.superess', *csője* 'pipe.3sgposs', *csőnk* 'pipe.1plposs', *csőtök* 'pipe.2plposs', *csőjük* 'pipe.3plposs' mainly refer to the ear of corn. – The noun *mű* 'work (of art)' from type 8C9 takes the *-v-* in front of all suffixes preceded by a linking vowel, e.g., *művön* 'work.superess', *művet* 'work.acc', *műve* 'work.3sgposs'. In many of its compounds, however, the use of the simple stem form is on the increase, like in the case of the 8C2 type *óramű* 'clockwork', e.g., *óraművön* 'clockwork.superess' vs. *óraműn* 'clockwork.superess', *óraművet* 'clockwork.acc' vs. *óraműt* 'clockwork.acc' and *óraműve* 'clockwork.3sgposs' vs. *óraműje* 'clockwork.3sgposs', etc. The posterior constituent of nine such compounds refers to some kind of mechanism or industrial institution.

A special sense distinction has developed in the inflected forms of the noun *hő* 'heat'. In its first sense it refers to a concept of physics, and as such, its stem is unchanged throughout, e.g., *hője* 'heat.3sgposs' refers the thermal energy possessed by something. However, if it is used in its original sense to refer to hotness of temperature, the suffixes are attached to a different variety of the stem, like *hevestül* 'ardour.soc<sup>7</sup>', *heve* 'ardour.3sgposs', *hevük* 'ardour.1plposs', *heveim* 'ardour.1sgposs'. Moreover, this differentiation extends to the adjectival derivatives as well, e.g., *hőjű* 'having a certain temperature' is different from *hevű* 'having ardour'. (The split between *hő* 'heat' and *hév* 'ardour' is reminiscent of the relationship between *só* 'salt' and *sav* 'acid', but the inflected forms of the former pair have not reached a total split yet.)

Among the irregular nouns having several stems, the difference between the meaning of *anyja* 'mother.3sgposs' (the person giving birth to somebody or the woman belonging to an older generation substituting the birth mother) and *anyája* 'mother.3sgposs', which does not refer to a family relationship (*Budapest legjobb anyája* 'the best mother in Budapest') is to be mentioned here. The same paradigm applies to the 9A words *apa* 'father' and *atyja* 'father'. The only difference between their compound forms is that the plural possessive form of the word *anya* 'mother' used in the former sense is primarily *anyái* 'mother.3sgposs.pl' (the form *anyjai* 'mother.3sgposs.pl' being very rare, since everybody has only one birth mother), while compound forms of the type *édesanya* 'real mother' are more likely to be used in the form *édesanyjai* 'real mother.3sgposs.pl'. The noun *ősanya* 'ancestress' (together

<sup>7</sup> *Soc* denotes the sociative suffix.



with *ősapa* 'ancestor' and *szépanya* 'great-great-grandmother') does not necessarily refer to one person, therefore only in the singular do we have *ősanyja* 'ancestress.3sgposs' and *ősanyjuk* 'ancestress.3pl.poss' (together with the less frequent form *ősanyája* 'ancestress.3sgposs', which does not refer to a family relationship), in the plural only the forms *ősanyáim* 'ancestress.1sgposs.pl', *ősanyái* 'ancestress.3sgposs.pl', etc. are used. In the type represented by *városatya* 'alderman', this order is reversed: the forms *városatyája* 'alderman.3sgposs' and *városatyájuk* 'alderman.3pl.poss' is the primary variant. – The noun *ifjú* 'young man' (together with *férfi* 'man' and *varjú* 'crow') from category 9A1, in addition to its alternation between the forms *ifja* 'young man.3sgposs' and *ifjúja* 'young man.3sgposs', can appear with several stems in the accusative case as well: e.g., *ifjút* 'young man.acc' (*ifjat* 'young man.acc'), and in the plural *ifjak* 'young man.pl' (*ifjúk* 'young man.pl'). The noun *gyapjú* 'wool', which is inflected in a similar way, can express two different kinds of relationships together with a personal possessive suffix: the wool belonging to the sheep's body is referred to by the form *gyapja* 'wool.3sg.poss', but the wool owned by a person is referred to by the form *gyapjúja* 'wool.3sgposs'. – The inflected forms of the (9A2) noun *fiú* 'boy' expressing a family relationship ('son') are *fiul* 'as a son' (*fiul fogad valakit* 'adopt somebody as a son'), *fiat* 'son.acc' (in *fiat szül* 'give birth to a son'), *fiak* 'son.pl' (meaning 'descendants'), *fiastul* 'son.soc', *fiam* 'son.1sgposs', *fiunk* 'son.1pl.poss', *fiuk* 'son.3pl.poss', but the possessive forms expressing other kinds of relationship are *fiúul* 'as a boy', *fiút* 'boy.acc', *fiúk* 'boy.pl', *fiústul* 'boy.soc', *fiúm* 'boy.1sgposs', *fiúnk* 'boy.1pl.poss', *fiújuk* 'boy.3pl.poss'. Among its compound forms only the ones which can express family relationships manifest a differentiation illustrated below, e.g., *gyámfiú* 'foster son', *keresztfiú* 'godson', *kisfiú* 'small boy or son', while the stems of the other compound words ending in *-fiú* are unchanged.

The form *fia* 'son.3sgposs' can also be regarded as an inflected form of the (by now) rare noun *fi* 'boy', expressing a family relationship, meaning 'the son of somebody'. However, the form *fija* 'man.3sgposs' is also possible, referring to a man originating from somewhere. This latter form bearing a fuller personal inflection is more characteristic of the compounds ending in *-fi*.

The distinction between the senses of the noun *nő* 'woman' through different personal possessive suffixes has almost lead to a complete split of meaning: the possessive form *nője* 'woman.3sgposs' in the nineteenth century primarily referred to a person's wife, but today this meaning is rather expressed by the form *neje* 'woman.3sgposs', while the form *nője* denotes a person's lover, girlfriend or (perhaps) female employee. The forms *nőm* 'woman.1sgposs',

*nőd* 'woman.2sgposs', *nőnk* 'woman.1plposs', *nőtök* 'woman.2plposs', *nőjük* 'woman.3plposs' used to refer to my, your, his, etc. wife, but today it refers to any woman other than one's wife. The marital relationship is expressed with the stem variant, which spread from the third person form to the others: *nejem* 'wife.1sgposs', *nejed* 'wife.2sgposs', *nejünk* 'wife.1plposs', etc.'. This split of forms will lead to a complete word split if the stem *nej*, at the moment used only for a humorous effect, will take its place in the repository of lexemes.

Among the pronouns referring to quantities, a sense distinction can be observed between the forms bearing a personal possessive suffix: the forms *annyink* 'that much.1plposs', *annyitok* 'that much.2plposs' and *annyiuk* 'that much.3plposs' are used to express that there is a certain amount of something in our, your or their possession. The forms *annyiunk* 'that many.1plposs', *annyiatok* 'that many.2plposs' or *annyiotok* 'that many.2plposs' and *annyiuk* 'that many.3plposs' refer to the number of persons (in a partitive sense): that many of us, you, them. The compound forms of *annyi* 'that much' and its front-vowel variants also show this duality: *ennyink* 'this much.1plposs' *van* (we have this much), *ennyitek* 'this much.2plposs' *van* (you have this much), *ennyiük* 'this much.3plposs' *van* (they have this much), but *ennyiünk* 'this many.1plposs', *ennyietek* 'this many.2plposs' (or: *ennyiőtök* 'this many.2plposs') and *ennyiük* 'this many.3plposs' *volt jelen* ('this many of us, you, them was present').

In the nominal use of ordinal numbers, which is characterised by an incomplete personal inflectional paradigm, the third person form without *-j-* has a special meaning, since it refers to a certain day of the month e.g., the (22A2) *annyiadika* 'that.3sgposs' (meaning: 'that day of the month'), *hányadika* 'which.3sgposs', *harmadika* 'third.3sgposs', *harmincadika* 'thirtieth.3sg poss', *hatodika* 'sixth.3sgposs', *huszadika* 'twentieth.3sgposs', *másodika* 'second.3sgposs', *nyolcadika* 'eighth.3sgposs', *sokadika* 'many.3sgposs' (meaning 'a long time since payday'); (22B2) *hetedike* 'seventh.3sgposs', *kilencedike* 'ninth.3sgposs', *negyedike* 'fourth.3sgposs', *ötödike* 'fifth.3sgposs', *tizedike* 'tenth.3sgposs', *tizenegyedike* 'eleventh.3sgposs' and *tizenkettedike* 'twelfth.3sgposs'. As opposed to this, the forms *annyiadikja* 'that.3sgposs', *hányadikja* 'which.3sgposs', *harmadikja* 'third.3sgposs', *hetedikje* 'seventh.3sgposs' are only occasionally used as nouns, e.g., *Két harmadik osztály van az iskolánkban, a kollégám harmadikja tehetségesebb* 'There are two third forms in our school, but the third form belonging to my colleague is more talented'. Some cardinal numbers also show a certain amount of sense distinction in this kind of use: cf. (22A4) *hatvana* 'sixty.3sgposs' *van* ('he has 60 points in a certain game, in cards, or 60 forints') and *hatvanja* 'sixty.3sgposs' *van* ('he has 60 out of some other, previously mentioned things'). The same distinction exists between *hetvene* 'seventy.3sgposs'

and *hetvenje* 'seventy.3sgposs' as well. The suffix *-uk/ük*, used to refer to several possessors, has a distributive sense, and refers to people: the forms *hatvanuk* 'sixty.3plposs', *hetvenük* 'seventy.3plposs' mean 'sixty or seventy people of them'. This sense is sometimes present even in singular form, e.g., *Nagy tömeg gyűlt össze, hetvenét különválasztották* 'A big crowd gathered, 70 of them were separated'. The two senses referred to above can be distinguished in the second person plural form as well: *könyvből hetvenetek* 'seventy.2plposs' (or *hetvenetek* 'seventy.2plposs') *van* ('of books, you have seventy'), but *hetveneteket* 'seventy.2plposs' *kiválasztottak* ('seventy people have been chosen from among you'). The corresponding forms of the (23A12) *hat* 'six' are the following: *hata* 'six.3sgposs' ('six out of it') vs. *hatuk* 'six.3plposs' and *hatja* 'six.3sgposs' ('six things in his possession') vs. *hatjuk* 'six.3plposs'. The forms for *hét* 'seven', from 29B are *hete* 'seven.3sgposs' ('seven out of it') and *hétje* 'seven.3sgposs' ('seven things in his possession'). For *öt* 'five', from category 23C5: *öte* 'five.3sgposs' ('five from it') and *ötük* 'five.3plposs' vs. *ötje* 'five.3sgposs' and *ötjük* 'five.3plposs'. The following forms of the nominal *egy* 'one' also carry different meanings: *együnk* [ɛj:yŋk] 'one.1plposs', *egyetek* [ɛj:ɛtek] 'one.2plposs', mean that we, you, etc. have one of something and the (rare or dialectal) forms *egyőnk* [ɛj:ø:ŋk] 'one.1plposs', *egyitek* [ɛj:itek] 'one.2plposs' or *egyőtök* [ɛj:ø:tök] 'one.2plposs' mean that one of us, you, etc. is referred to. Similarly, the form *négyünk* 'four.1plposs' and *négyetek* 'four.2plposs', *négyük* 'four.3plposs' can be used in any sense, but *négyőnk* 'four.1plposs', *négyőtök* 'four.2plposs' and *négyőjük* 'four.3plposs' is only used to refer to four people out of us, you, etc. *Többetek* (or *többötök*) 'more.2plposs' refers to the fact that you have several of something, but *többőtök* 'more.2plposs' means 'several of you'. The (22D) forms *sokja* 'much.3sgposs', *sokatok* (or *sokotok*) 'much.2plposs' and *sokjuk* 'much.3plposs' in *sokja/sokatok/sokjuk van* 'there is much in his/your/their possession' refer to people's wealth, while the form *soka* 'many.3sgposs' in *az esetek, emberek soka* refers to 'many of the cases, the people', *sokatok* or *soktok* (rarely *sokótok*) 'many.2plposs' and *sokuk* 'many.3plposs' (more usual) mean 'a lot of you, or them'. The (26C1) form *kettője* 'two.3sgposs' refers to two pieces, objects or persons, but *ketteje* 'two.3sgposs' means 'both of the people, or two out of them'. *Kettőjük* 'two.3plposs' can be used in any sense, but *kettejük* 'two.3plposs' means 'two of them, or they both'. – Among the numerals with alternating stems we find forms with a characteristic partitive personal suffix in the group of the (29A) *néhány* 'some' and (29A1) *három* 'three': the forms *hányunk* 'how many.1plposs', *néhányunk* 'some.1plposs', *hányotok* 'how many.1plposs', *néhányotok* 'some.2plposs' (or *hányatok* 'how many.2plposs',

*néhányatok* 'some.2plposs'), *hányuk* 'how many.3plposs', *néhányuk* 'some.3plposs' can be used to express any kind of possession, but the forms *hányónk* 'how many.1plposs', *néhányónk* 'some.1plposs', *hányótok* 'how many.2plposs', *néhányótok* 'some.2plposs', *hányójuk* 'how many.3plposs' and *néhányójuk* 'some.3plposs' can only be used in the sense of 'some of them, etc.', in a very restricted way, in addition to some of their compound forms. Beside the forms *hármunk* 'three.1plposs', *hármatok* 'three.2plposs', *hármuk* 'three.3plposs', the forms *hármónk* 'three.1plposs', *hármótok* 'three.2plposs' and *hármójuk* 'three.3plposs' in the sense of 'three of us, you or them' can also appear in restricted use.

The meaning of the personal possessive suffix has blurred so much in a couple of words that it gave rise to the mixing of certain grammatical and lexical functions. Such is the (26A2) *orja* 'spare-rib (of a pig)', which originally contained a personal possessive suffix (already blurred). Those who consider it a root form, add further personal suffixes to it, e.g., *orjám* 'spare-rib.1sgposs', *orjád* 'spare-rib.2sgposs', *orjája* 'spare-rib.3sgposs', etc., but the latter forms can only refer to a type of pork. As the name of a body part of a pig, however *-ja* does have the value of a personal suffix, e.g., *a sertés orja* 'the spare-rib of the pig'. The (26B2) word *fehérje* 'egg white' also behaves the same way. In its original sense it contained a personal suffix, e.g., *a tojás fehérje (és sárgája)* 'the white part of the egg (and its yolk)'. In a chemical and biological sense the word *fehérje* 'protein' is a root word, therefore further personal possessive suffixes are added to it, e.g., *fehérjém* 'protein.1sgposs', *fehérjéje* 'protein.3sgposs', etc. – Forms of other words containing a personal possessive suffix, when used in addresses, can acquire a pronominal-like meaning, and in some of their suffixed forms the functions of naming and addressing are also separated. For example, the form *őnagysága* 'madam' was both used to refer to and to address women (according to certain social conventions), while the plural forms show a sense distinction, since *őnagyságaik* was only used for referring to, while *őnagyságáik* for addressing women. The (28A4) word *nagysád* 'madam' shows a sense differentiation in several of its forms: as a (pronominal) form of address, the forms *nagysád* 'your ladyship' (referring to the second person singular, as indicated by the suffix) and *nagysádtok* 'your ladyship' (the plural of the pronominal form) is used, for naming somebody in third person, we use the same form *nagysád* in the singular, with the *-d* affix incorporated into the stem: The plural form of the word used in the latter sense is *nagysádok* 'ladies', and it can also take further personal possessive suffixes, e.g., *nagysádom* 'lady.1sgposs', *nagysádod* 'lady.2sgposs' and *nagysádja* 'lady.3sgposs'. It has an idiosyncratic plural form, *nagysádek* (which can ei-

ther be used to address more ladies and to refer to a company consisting of a lady and some other people).

The fact that a sense distinction has developed between various suffixed forms of the noun *nő* 'woman' and that the form *neje* is in the process of becoming a lexeme is also caused by the blurring of the meaning of the personal possessive suffix. This latter feature was also mentioned above in connection with the sense distinctions of kinship terms.

7. As far as the idiosyncratic forms of adjectives are concerned, the inflectional affixes of the manner adverbial, *-lag*, *-an* and *-ul*, which look like derivational affixes, are not phonological variants of each other. Instead, they are alternants which can be substituted freely for each other in certain paradigms, in certain others, only one of them is allowed depending on the phonological characteristics of the stem, while in some others they can be used to distinguish between different senses of the word. These differences, however, are not characteristic of the meaning of the stem, but rather of the syntactic role of the adverbial form. Among alternating suffixes, the form *-lag/-leg* is the most similar to derivational suffixes (since a further derivational affix can also be attached to it), and in some cases it can also lead to the formation of a lexicalised adverb, while the form with the suffix *-an/-en* has retained more of the original participial feature of the adjective. The above suffixes can be attached to the adjective (11A2) *tagadó* 'denying': the form *tagadóan* means 'negating', while *tagadólag* 'in a way expressing denial', e.g., *tagadólag int* is 'nod one's refusal'. The adverbs formed from the corresponding word *állító* 'affirmative' show a greater difference of meaning: *állítóan* means 'formulating something in an assertive way', while the form *állítólag* is a separate lexical entry, meaning 'according to the assertion of somebody else'. In the adjectives with a participial origin like *bíráló* 'criticising', *csillapító* 'relieving', *pergátló* 'preventing legal action', *vonatkozó* 'referring', there is hardly any difference between the meanings of the forms with the two types of adverbial affixes. However, attaching the adverbial suffixes to the words *bűnbánó* 'repenting' or *vádló* 'accusatory' results in forms with a slight meaning distinction, cf. *bűnbánóan* 'repenting his sins', *vádlóan* 'accusing' vs. *bűnbánólag* 'showing (or perhaps only pretending) repentance' and *vádlólag* 'expressing accusation'. From the adjectives *felszabadító* 'liberating', *tanító* 'teaching' and *tisztító* 'cleaning' we can form *felszabadítóan*, *tanítóan* and *tisztítóan* to express a concrete meaning (if these forms can be used at all), while the forms *felszabadítólag*, *tanítólag*, *tisztítólag* refer to an emotional or intellectual influence, meaning 'causing a mental relief', 'causing edification' and 'causing (mental, spiritual) purification'. The

adverb *folyóan* ‘in a flowing manner’ also has a concrete meaning, while the form *folyólag* has a more abstract sense, meaning ‘as a result of something’. *Hívogatóan* and *szárítóan* only mean ‘inviting, in an inviting manner’, and ‘drying, in a drying manner’, while the result of inviting and the possibility of allurements is suggested by the form *hívogatólag*, and the expected success of drying by the form *szárítólag*. The same difference can be observed between the concrete meaning of *irányítóan* ‘in a directing manner’, *javítóan* ‘in a repairing manner’, *óvóan* ‘in a protecting manner’ and *pergátlóan* ‘in a manner preventing legal action’ and the more abstract senses of *irányítólag*, *javítólag*, *óvólag*. While the form *látszóan* ‘in a way that it seems’ is used in a more concrete sense, based on the original meaning of the participle, the meaning of the word *látszólag* ‘according to appearance, but not in reality’ has become lexicalised and isolated from it. The word *kizáróan* is simply ‘in an exclusive manner, excluding’, while the form *kizárólag* has turned into an adverb with a specific meaning ‘only (i.e., excluding everything else)’.

The inflected forms of the word *cselekvő* ‘acting’ (containing front vowels), namely *cselekvően* ‘in an acting manner’, which is used in a sense corresponding to the original meaning of the participle and the adjective-based *cselekvőleg* ‘actively’ only show a slight difference in meaning, but among the inflected forms of *ellenkező* ‘opposing’ only *ellenkezően* expresses real opposition, while *ellenkezőleg* means ‘the other way round, in the opposite case’. The forms *fékezően* ‘in a braking manner’, *megrendítően* ‘in a moving manner’ and *ösztönzően* ‘in a stimulating way’ are participial adverbials, while *fékezőleg*, *megrendítőleg* and *ösztönzőleg* are adjectival adverbials, which express the success of the influence referred to by the adjective. There is, however, no significant difference between the generally more abstract forms *felemelően* ‘in an elevating manner’, *feltehetően* ‘supposedly’, *kimerítően* ‘exhaustively’ and *számtottevően* ‘significantly’ and *felemelőleg*, *feltehetőleg*, *kimerítőleg* and *számtottevőleg*. The form *kártevően* ‘in a harmful manner’ refers more to some human, while the form *kártevőleg* more to some physical or chemical action. The form *kötelezően* ‘in a compulsory manner’ can be used for everything and everybody, who or what obliges, while the form *kötelezőleg* can only refer to a law or an order. Such a difference also exists between *perdöntően* ‘in a decisive manner’ and *perdöntőleg*. The forms *lenézően* and *lenézőleg* are both used in an abstract sense, meaning ‘in a despising manner’, but the form with the suffix *-leg* refers more to a facial expression. The difference is the same between the forms *megvetően* ‘contemptuously’ and *megvetőleg* as well. – The distinction between *remélhetően* ‘hopefully’ and *remélhetőleg* reminds one of the difference between *állítóan* vs. *állítólag*, especially since the adverb *remélhetőleg* as

a sentence adverb also expresses modality. This means that it restricts the force of the following statement, and it can also be turned into a main clause predicate, e.g., *remélhető, hogy...* 'it is to be hoped that...'. The same difference can be observed between the forms *sejthetően* and *sejthetőleg* 'in a guessable manner'. The adjective *sürgető* 'hurrying' can either be turned into *sürgetően* 'in a hurry' or *sürgetőleg*, but the latter one can only be regarded the adverbial form of the adjective meaning 'in a hurrying manner'. – The same behaviour characterises the following adjectives with alternating forms in the plural, like the (17A2) *látható* 'visible', since *láthatóan* (*láthatón*) means 'in a visible manner, perceptible with the eyes', while *láthatólag* means 'in a conceivable manner'; the (17C2) *eltérő* 'different', where the form *eltérően* means 'in a different manner', while *eltérőleg* is 'contrary to' (e.g., *eltérőleg nyilatkozik* 'express a contrary opinion'). *Feddően* means 'in a reproving manner', while *feddőleg* is 'with the intention to reprove'; *fürkészően* is 'searchingly', while *fürkészőleg* is 'with the intention to search'; *hihetően* is 'in a credible manner' (manner adverb), while *hihetőleg* is 'presumably' (a sentence adverb, having modal value); the (17C5) *előzően* is 'overtaking' (as a participle), while *előzőleg* means 'before the thing mentioned' (as a time adverbial); *illetően* means 'in a manner relating to somebody', while *illetőleg* is 'concerning somebody' (also as a conjunction:) 'respectively', *visszamenően* is 'going back' (also as a participle), while *visszamenőleg* is 'going backwards in time' (as an adverb); the (17C13) *következően* means 'in a following manner', *következőképp(en)* is 'as follows', while *következőleg* is 'as a result'. – Among the members of group 18, which always form their plural with the help of a linking vowel, the following distinctions can be observed: *atyai* 'fatherly' is inflected as *atyaian* if it is to mean 'in a manner characteristic of a father', while as *atyailag* if it is to mean 'as a father'; *alanyian* means 'in an individual way', while the form *alanyilag* 'as an individual' occurs only in legal usage; *bajtársian* means 'as a comrade', while *bajtársilag* 'from the point of view of comrades'; *barátian* is 'in a friendly manner', while *barátilag* means 'as a friend'; *drákóian* means 'very severely', while *drákóilag* is 'in a manner characteristic of Dracon'; *drámaian* means 'in a dramatic way', but *drámailag* is 'with the means of drama'; *egyhangúan* is 'without change', while *egyhangúlag* means 'unanimously'; *egyoldalúan* means 'in a one-sided manner', while *egyoldalúlag* is 'without the other party' (in a legal sense); *gyakorlatian* means 'in a practical manner', while *gyakorlatilag* is 'in practice'; *irodalmian* is 'in a literary manner', while *irodalmilag* means 'from the literary point of view'; *királyian* is 'in a kingly manner', while *királyilag* is 'originating from the king'; *líraian* is 'in a manner expressing emotions', while *lírailag* is 'from the point of view of poetry'; *polgárian* is 'in a civic manner',

while *polgárilag* 'in a civil case'; *prózaian* means 'in an everyday manner', while *prózailag* means 'in prose'; *salamonian* is 'in a wise and just manner', while *salamonilag* is 'in a manner characteristic of King Solomon'; *szónokian* is 'in a manner characteristic of orators', while *szónokilag* means 'from the point of view of oratorship'; *zsarnokian* is 'in a tyrannic manner', while *zsarnokilag* is 'as a tyrant'. The same differences can be observed in other paradigm types as well: the adjective *kétoldalú* usually takes the *-an* suffix, while the form *kétoldalúlag* means 'applying to both parties'; *zárdaiian* is 'in a manner characteristic of a convent', but *zárdailag* is 'through the cloister'. Corresponding forms with front vowels are *egyéniien* 'in an individual, characteristic way' vs. *egyénilag* 'as an individual'; *emberien* 'in a manner suiting a person' vs. *emberileg* 'from a human point of view'; *eredetien* 'in an individual manner' vs. *eredetileg* 'in its original state' or 'concerning its origin'; *érzékien* 'in a way appealing to sexuality' vs. *érzékileg* 'in a perceivable way'; *eszményien* 'wonderfully' vs. *eszményileg* 'in theory'; *fejedelmiien* 'magnificently' vs. *fejedelmileg* 'originating from the emperor'; *festőien* 'in a colorful manner' vs. *festőileg* 'from the painter's point of view'; *gyermekien* 'in a manner characteristic of children' vs. *gyermekileg* 'as a child'; *költőien* 'rich in emotions' vs. *költőileg* 'as a poet'; *művésziien* 'in a skillful manner' vs. *művésziileg* 'from an artistic point of view'; *zeneien* 'in a manner resembling music' vs. *zeneileg* 'with music, from the point of view of music'; *hercegiien* 'in a princely manner' vs. *hercegileg* 'originating from a prince'; *színésziien* 'in a theatrical way' vs. *színésziileg* 'from the point of view of acting'; *szakértőien* 'like an expert' and *szakértőileg* 'as an expert'.

The inflected adverbial forms of the (12A1) *vad* meaning either 'wild' or 'fierce' show an interesting distribution. The form *vadul* 'fiercely' is a manner adverbial (e.g., *vadul támad* 'attacks fiercely'), while the forms *vadon* 'wild' and *vadan* 'in a fierce manner' are state adverbials showing a slight difference of meaning, cf. *vadon tenyészik* 'grow wild' (a permanent state) vs. *ilyen vadan nem maradhatsz köztünk* 'as long as you are so fierce, you cannot stay with us' (a temporary, changeable state). – The adjective *igaz* 'true' also shows some idiosyncratic characteristics. Beside the form *igazán* 'really, very much', which has already turned into an adverb, there are the forms *igazul* 'in a true manner' and *igazan* 'in a state of being true' (e.g., *igazan küzdötte végig a háborút* 'he was true while he fought till the end of the war').

In most cases, only the suffix *-ul/ül* of the manner adverbial and the suffix *-an/en* of the state adverbial show an opposition of meaning, e.g., *absztrak-tul fogalmaz* 'speaks in an abstract manner' vs. *absztraktan marad* 'stay in an abstract state'; *egzaktul gondolkozik* 'think in an exact manner' vs. *egzaktan te-kintve* 'considering it objectively'; *flottul él* 'live freely, without inhibitions' vs.



*flottan érzi magát* 'feel quite well', *flottan jelenik meg* (e.g. *a kiránduláson*) 'appear somewhere (e.g. at an excursion) light, with little luggage'; *nyomorultul elpusztul* 'die a miserable death' vs. *nyomorultan él* 'live wretchedly'; *szilárdul ragaszkodik valamihez* 'insist firmly on something' vs. *szilárdan marad* 'stay firm'; (12B3) *remekül* 'in a perfect state or manner' vs. *remeken* 'in a perfect state'. The adjectives in category 12B4 refer to a state or manner if they take the *-en* suffix and only to a manner if they take the *-ül* suffix, e.g., *dekadens* 'decadent', *diszkrét* 'discrete', *extrém* 'extreme' and 20 other words of Latin (occasionally French or German) origin, in addition to *merev* 'stiff' and *rest* 'lazy'. Some adjectives like *megveszekedett* 'cursed' and especially *veszett* 'mad' can also be used as a degree adverbial with an *-ül* suffix, like *veszettül csinos* 'extremely pretty'. The sense distinction is sometimes not very clear-cut. But there exist some adverbs ending in the suffix *-ul* with a meaning completely isolated from the original meaning of the adjective (or participle), e.g., *átkozottul* 'in a cursed manner, very much'; *rohadtul* 'nastily' (used for behaviour). The word *finoman* 'finely' or 'gently' can either be a manner or a state adverbial, but the variant *finomul* contains an idiosyncratic shade of meaning corresponding to 'minutely' or, used in a sarcastic manner, 'cunningly'.

The suffix *-ul/ül* is more often connected to figurative meanings. The separation of the state and manner adverbials is especially characteristic of type 12A4, e.g. *csúf* 'ugly', *falánk* 'greedy', *hazárd* 'hazardous', *intakt* 'intact', *konzervatív* 'conservative', *morc* 'morose', *naiiv* 'naive', *nyápic* 'weakling', *rút* 'ugly', *torz* 'deformed', *undok* 'disgusting', *virgonc* 'lively', etc.

The numeral *egy* 'one' served as the base form for some idiosyncratic formations like the adverbials *egyként*, *egyképp* or *egyképpen* 'in the same way'. The state adverbial *egyén* [ej:ɛn] 'one of them' is only rarely used, the form is rather characteristic of compounds ending in *-egy*, e.g., *tizenegyen voltak* 'there were eleven of them'.

The suffix *-képpen* can only refer to manner, while *-an/-en* can also refer to a state, e.g., (17A6) *hasonlóan jelent meg* 'it appeared with a similar look', but *hasonlóképp(en) gondolkodik* 'he thinks in a similar way'; (17A9) *méltó(a)n* 'in a manner worthy of something' (e.g., *művelt emberhez méltóan* 'in a manner worthy of an educated person') but *méltóképpen* 'to an extent to which he deserves on the basis of his importance'.

The (14A2) *színjőzan* 'perfectly sober' is more of a manner adverbial if it takes a suffix *-ul* (*színjőzanul él* 'he lives a sober life'), and it is more of a state adverbial if it takes a suffix *-on* or *-an* (*színjőzanon* or *színjőzanan jött be* 'he came in sober'). There is a distinction similar to this latter one between the forms of the (14A8) *dévaj* 'frolic', *duhaj* 'rowdy', *fanyar* 'sour',

*monoton* 'monotone' which take the same suffixes. The same kind of difference is observed between the forms *józanul* and *józanan* meaning 'sober', while the adjective *pazar* 'lavish', which is inflected in the same way, acts as a manner adverbial. The corresponding forms of *pimasz* 'impertinent', namely *pimaszul* 'in an impertinent manner' vs. *pimaszan* 'in an impertinent state' show a distinction, which is not observable in the forms of *botor* 'stupid', *fukar* 'mean' and *jámbor* 'pious'. – Among the two corresponding forms of the (15B6) *fösvény* 'mean', namely *fösvényen* and *fösvényül*, the former is perhaps more of a state adverbial, while the latter is more of a manner adverbial, although the difference is almost as much blurred as that between the two forms of *fukar*. As opposed to this, both of the suffixed forms of the (14B7) *meztelen* 'naked', namely *meztelenül*, or the rarer *meztelenen* can only be taken as state adverbials, since the original adjective also refers to a temporary state. – Even the suffixed forms of the (14D2) *kóser* 'kosher' and *sóher* 'mean', containing both front and back vowels show such a distinction: the forms with *-ül(-ul)* are more like manner adverbials, while those with *-en* are more like state adverbials. The difference is similar between the (15B6) forms *lezseren* 'in a relaxed manner' and *lezserül*.

The above distinctions illustrated above also show up in type 16 with alternating stem-final vowels: the adverb (16A3) *csúnyán* 'in an ugly manner' can either be a manner or a state adverbial (*csúnyán rászed valakit* 'deceive somebody badly' and *ne menj ki ilyen csúnyán* 'don't go out while you are so ugly'), and as an adverbial of result as well: *csúnyán ír* 'he writes in an ugly way', while the form *csúnyául* 'in an ugly manner' can only be a manner adverbial (*csúnyául rászedte* 'he deceived him badly'). The above division of functions is seen in the manner and state adverbial forms of the following adverbs expressing qualification, e.g., *balga* 'silly', *bamba* 'foolish', (e.g., *bambául nevet* 'laugh in a foolish way'), *buta* 'stupid' (*bután néz* 'look in a stupid way', but *bután* or *butául viselkedik* 'behave in a stupid way') and half a dozen other adjectives, but the forms with *-ul*, some of which are characteristic of regional dialects, are used only by a few people. The difference is more clear-cut in the adjectival sense of the word *marha* (cattle): *nagyon marhán válaszolt* 'he answered in a foolish way' (manner adverb) vs. *marhául lesziadták* 'they scolded him very much' (degree adverbial).

Some characteristic sense distinctions from group 19, which displays irregular alternations of the end of the stem, are the following: *hosszan* is 'long, long-lasting' but *hosszúan* is 'in a long state'; *éberen* is 'watching keenly, ready to wake up', while *ébren* means 'not sleeping'; *híven* is 'in a faithful manner' while *hűen* (*hűn*) means 'in a state of being faithful'; *könnyen* is 'without any

difficulty' while *könnyűen* is 'in a light state'. – The (19C8) *különb* 'better' is an idiosyncratic form developed from the comparative form of an adjective (*különbül* 'in a better way' is a manner adverbial, while *különben* 'otherwise' is a sentence adverbial).

8. Only in one case does a variant of the comparative suffix give rise to a separate meaning: the regular comparative form of the (15C6) *idős* is *idősebb*, while the archaic form *idősb* refers to the elder of two people with the same names.

The superlative form of adjectives referring to spatial location is formed with the *leg-* prefix but without the *-bb* suffix, e.g., *alsó* 'lower', *hátsó* 'rear', *utolsó* 'last', *túlsó* 'opposite', *elülső* 'front', *felső* 'upper', *belső* 'inner', *végso* 'last'; it can, however, receive a suffix *-bb* as well, when it is used in a figurative sense, e.g., *legalsó* or *legalsóbb* 'the one with the lowest rank', *legfelsőbb* (*legfelső*) 'the one with the highest rank', *legszélsőbb* 'the most radical one', *legbensőbb* (*legbenső*) 'the most confidential one', and *a legvégső(bb) esetben* 'as a last resort'.

Some adjectives can have a comparative form which is produced by adding the derivative suffix *-i* to the regular comparative affix, which results in a slight change of meaning, e.g., the two forms of the (10B10) *közeli* 'close' (*közelebbi* is 'closer in time, space or any other sense vs. *közelibb* 'closer only in space'); and those of the (19B11) *régi* 'old' (*régibb* is 'existing for a longer time' and *régebbi* is 'earlier').<sup>8</sup>

9. The plural suffix can attach to adjectives ending in *-ó* either in the form of *-k* or *-ak*. In order to investigate the possibilities of sense division, we have to take into account the tendency observed by Éva Ruzsiczky in the fifties (see Ruzsiczky 1955), according to which the adjectives ending in the derivational suffixes *-ó* and *-ő* are in the process of separating from the corresponding nouns in the plural form due to their propensity to attach the plural suffix to the stem with a linking vowel. The above tendency, however, does not extend to adjectival participles, therefore the following tripartition has developed: the plural form of nouns ends in *-ók*, the plural suffix of adjectives is *-óak* or *-ók*, while the plural suffix of adjectival participles is *-ók*. (In the last decades the

<sup>8</sup> See also Hexendorf (1977).

latter distinction has somewhat blurred, since some people have started to use the plural form *-óak* for words with a participial meaning.)<sup>9</sup>

The above differentiation according to parts of speech is connected to some of the individual sense differentiations found with adjectives in category 17. Here we also find the following distinction. While the plural form with a linking vowel expresses a qualification (mainly in the abstract sense), form with the bare *-k* refers to a classification into a category (in a more concrete, stricter sense in most cases). Certain looks, sounds, smells can be referred to by the forms *áthatóak* or *áthatók* 'penetrating.pl', but for the property of transitive verbs only the form *áthatók* is possible. The forms *beláthatóak* and *beláthatók* 'conceivable.pl' refer to a property of statements or truths, while *beláthatók* alone can denote a property of distances, namely, being 'discernible'. When people can understand and accept other people's actions, moreover, their faults as well, they can be referred to by the word *belátóak* 'considerate.pl', while those who can see into or through some, are denoted by the word *belátók* (mainly in a participial or temporally nominalised sense). Arguments can be decisive, which is expressed by the forms *bizonyítóak* or *bizonyítók* 'decisive', while those people who prove something can only be referred to by *bizonyítók* 'verifier.pl'. The form *hódítóak* can only denote people who excite attraction or admiration, in other senses the form *hódítók* 'conqueror.pl' is appropriate.

10. Among adverbials bearing personal suffixes (or, rather: personal pronominal adverbials) only one irregular paradigm is used to express a special meaning. In addition to the *nálam, nálad, nála, nálunk, nálatok, náluk* 'at me/you/him/us/you/them', also 'than me/you/him/us/you/them' forms in general use, the paradigm containing a repetition of elements like *nálamnál* 'than me' can only appear with a comparative form of an adjective with a comparative meaning, mainly in the non-standard variety (e.g., *szébb vagyok, jobb vagyok náladnál* 'I'm prettier, better than you').

<sup>9</sup> The most reliable way to decide whether a certain form is used as an adjective or a participle is that the former sense can be questioned with the help of the question word *milyenek* 'what kind.pl' (meaning 'what kind are they?'), but the latter sense cannot.

## References

- Elekfi, L. 1996. Részleges szóhasadások mint határesetek. [Partial word splits as borderline cases]. In: Magyar Nyelv 151: 286–96.
- Hexendorf, E. 1977. A *régebbi* : *régibb* típusú alakpárokról [On the pairs of forms of the type *régebbi* : *régibb* 'older']. In: A magyar nyelv grammatikája. A magyar nyelvészek III. nemzetközi kongresszusának előadásai. [The grammar of Hungarian. Talks given at the Third Congress of Hungarian Linguists.] Nyelvtudományi Értekezések 104, 399–407. Akadémiai Kiadó, Budapest.
- Ruzsiczky, É. 1955. Az *-ó*, *-ő* végű igenevek kötőhangzós (*-ak*, *-ek*) többesének a kérdéséhez [On the *-ak/-ek* plural of participles ending in *-ó/-ő*]. In: Magyar Nyelvőr 79: 289–98.
- Tompa, J. 1957. A névszói kötőhangzó megkülönböztető szerepe [The distinguishing property of the linking vowel]. Nyelvtudományi Értekezések 14. Akadémiai Kiadó, Budapest.

Address of the author: László Elekfi  
 Research Institute for Linguistics  
 Hungarian Academy of Sciences  
 1399 Budapest  
 P. O. Box 701/518.  
 Hungary



## CASE MISMATCHES IN GREEK: EVIDENCE FOR THE AUTONOMY OF MORPHOLOGY\*

MANUEL ESPAÑOL-ECHEVARRÍA AND ANGELA RALLI

### Abstract

This paper investigates the Case properties of Clitic Left Dislocated (CLLD) constructions in Modern Greek. In particular, CLLD constructions involving Case mismatches. It is argued that specification for a  $+/-$ Genitive feature is involved in the morphological Case marking of Greek Nominatives [ $-$ Genitive], Accusatives [ $-$ Genitive], and Genitives [ $+$ Genitive], and that the syntactic operation of feature checking makes use of the [ $+/-$  Genitive] feature, instead of more fine-grained features for Case. This analysis leads to the conclusion that Case checking may involve partial (morphological) feature descriptions, and supports the view of morphology as a well-differentiated component of grammar.

### Introduction

This paper investigates the Case marking of dislocated constituents and resumptive clitics in Clitic Left Dislocated (CLLD) constructions in Modern Greek, a language in which Case is morphologically realized. We show that Case mismatches between the dislocated constituent and the clitic are possible, whereas mismatches in Gender, Person, or Number are not. We consider this evidence in the framework of Baker's (1996) **Chain Condition** relating clitics and dislocated constituents, and propose a reformulation of this condition in terms of the  $+/-$ Interpretable distinction proposed in Chomsky (1995). Furthermore, we investigate the conditions giving rise to Case mismatches in Greek CLLD constructions. From the consideration of CLLD free relative clauses, a typical context of Case mismatch, we conclude that these Case mismatches cannot be accounted for in a customary Minimalist treatment of Case. We claim that the syntactic operation of Case checking may involve partial feature

\* Preliminary versions of this paper have been presented at the Incontro di Grammatica Generativa (Verona, Feb. 1998) and at the International morphology Meeting (Budapest, June 1998). We thank the audiences of both meetings for their helpful remarks, Maria Luisa Rivero for written comments on an earlier draft of this paper, and Antonia Androutsopoulou for constructive discussion. All errors are our own.

descriptions. More concretely, we argue that specification for a +/–Genitive feature is involved in the morphological Case marking of Greek Nominatives [–Genitive], Accusatives [–Genitive], and Genitives [+Genitive], and that the syntactic operation of feature checking makes use of the [+/–Genitive] feature, instead of more fine-grained features for Case, also available in the morphological feature structure. This proposal has interesting consequences for the organization of grammar. First, it supports the view of morphology as a well-differentiated component of grammar. Second, it assigns a significant role to morphology in the grammar of Case, the most structurally relevant of all inflectional features (as defined by Kuryłowicz 1964). Finally, the present proposal claims that syntax may have access to the morphological feature structure of lexical items. The discussion is organized as follows. Section 1 contains the data relevant to our discussion, and deals with the theoretical consequences that the existence of Case mismatches in Greek CLLD contexts has for Baker's (1996) view of Clitic Left Dislocation. In section 2, we deal with the issue of feature matching in CLLD chains, cf. 2.1, and present our analysis of Case mismatch in CLLD contexts, cf. 2.2 and 2.3. Section 3 is devoted to analyse the sort of Case-checking operation leading to Case mismatches in CLLD free relatives as a case of syntactic syncretism. Some concluding remarks in section 4 close the paper.

## 1. Case mismatches in clitic left-dislocation constructions: the data

Elaborating on Cinque's (1990) analysis of Italian CLLD constructions, Baker (1996) claims that dislocated constituents in languages, both polysynthetic and non-polysynthetic, form a non-movement chain with a unique pronominal in argument position. The conditions on this chain formation are stated as follows (cf. Baker 1996, 112):

- (1) THE CHAIN CONDITION  
 X and Y may constitute a chain only if:
- (i) X c-commands Y.
  - (ii) X and Y are coindexed.
  - (iii) There is no barrier containing Y but not X
  - (iv) X and Y are nondistinct in morphosyntactic features (i.e., category, person, number, gender, case, etc.)

According to the condition in (1iv), feature mismatches between left-dislocated constituents and the pronominals coindexed with them are not allowed. Baker,



however, has observed many apparent cases of feature mismatch between the two constituents in polysynthetic languages such as Mohawk:

- (2) Sak wa-shukení-k^-' ne raó-skare' kanát-a-ku  
 Sak FACT-MsS/1DO-see-PUNC ne MsP-friend town-Ø-in  
 'Sak saw me with his girlfriend in town'  
 (Lit.: 'His girlfriend, Sak saw us two in town')<sup>1</sup>

In (2), the pronominal form coindexed with the object shows dual number, whereas the dislocated NP *his girlfriend* denotes a singular individual. The occurrence of feature mismatches in polysynthetic languages is explained by Baker (1996) as a consequence of the fact that, in these languages, DPs bear no morphosyntactic markers for Number, Gender, Person, or Case. Thus, coindexation between the adjoined DP and the VP-internal pronoun becomes possible, because condition (iv) in (1) is vacuously satisfied.

For a non-polysynthetic language like Greek, with overt markers for Number, Gender, Person, and Case in DPs and clitic pronouns, an account along these lines predicts that dislocated nominal constituents require matching between their morphosyntactic features and the features of the pronominal elements coindexed with them.<sup>2</sup> Any feature mismatch between a dislocated DP and the pronominal element (clitic) in its chain would lead to a violation of condition (iv) in (1), ruling out the CLLD construction. However, there are counterexamples to this prediction, particularly with respect to Case matching. Such examples have been discussed in Tzartanos (1946), Philippaki-Warburton-Stavrou (1986), Catsimali (1990), Philippaki-Warburton (1990), Tsimpli (1990), Alexiadou-Varlokosta (1996):

- (3) (a) Opjos m'agapai, ton agapo  
 whoever-nom me-acc-loves, him-acc love-1sg  
 'I love whoever loves me' (or *Whoever loves me, I love him*)
- (b) Opjon den grapsi sosta afto to thema, tha ton aporipso  
 whoever-acc not write-perf right this the topic fut him-acc discard-perf-1sg  
 'I will discard whoever does not write well this topic'

<sup>1</sup> From Baker (1996, 122), where "FACT" stands for *factual*, "PUNC" for *punctual*, "s" for *singular*, "S" for *subject*, "D" for *dual*, "M" for *masculine*, "O" for *object*, "1" for *first person*, and "P" for *possessor*.

<sup>2</sup> Inflectional class is another feature that overtly marks the inflectional system of Greek nominals, but, as discussed in Ralli (1997; to appear), it is syntactically irrelevant.

- (4) (?)I      fitites,                      i      kathigites                      tus                      agapane      olus<sup>3</sup>  
                  the students-nom, the professors-nom them-acc love-3pl all-acc  
                  ‘The professors love all the students’ (or *The students, the professors love them all*)

In (3a), the head of the dislocated free relative *opjos* occurs in nominative case, although it is coindexed with an accusative clitic, i.e., *ton*. In (3b), the head of the dislocated free relative *opjon* is in the same case (Accusative) as the coindexed clitic in the matrix clause, but this Accusative case cannot be checked against the verb of the relative which requires nominative case. In (4), the dislocated DP *i fitites* ‘the students’ shows nominative case, but it is coindexed with an accusative clitic, i.e., *tus* ‘them’. These Case mismatches pose a problem for Baker’s (1996) unified account of dislocated constituents in both polysynthetic and non-polysynthetic languages. In addition, the examples in (3) and (4) raise more general questions on Case theory. In (3b), the Nominative syntactic or abstract Case associated with the verb *grapsi* ‘write-perf’ does not show up on any lexical item of the sentence, whereas (3a) indicates that it may show up, cf. *opjos* ‘whoever-nom’. On the other hand, the nominative Case of the CLLD DP *i fitites* ‘the students-nom’ does not seem to be related to any syntactic Case assigned by the verb *agapane* ‘love-3pl’ in (4).

The examples in (3)–(4) call for a reformulation of Baker’s Chain Condition in (1), particularly, in relation to the nondistinctiveness in Case condition on the elements of a non-movement chain. This reformulation should be able to account also for the two following facts. (a) The relative pronoun or the DP must bear the Case marker checked by the matrix verb if they appear in a non-dislocated position. Compare (5a) to (3a) and (6a) to (4), respectively:

<sup>3</sup> Case mismatches such as the one illustrated in (4) are always slightly deviant, as indicated by the question mark, although grammatical. As noted by Tzartanos (1946), this type of Case mismatches are much better if some material (related to the dislocated constituent) intervenes between the dislocated constituent and the rest of the sentence:

- (i) I                      kira-Rini                      tu                      Kritu,                      tu                      Duka                      i  
                  the-nom                      miss-Irene-nom the-gen Kritos-gen, the-gen Duka-gen the-nom  
                  thigatera,                      xronia                      tis                      etimazun                      ta                      prikia  
                  daughter-nom, years-acc                      her-gen prepare-3pl the                      dot  
                  ‘Mrs. Irene of Kritos, Duka’s daughter, they prepare her dot for years’  
                  (from Tzartanos (1946, 264))

In spite of the difference in acceptability between (3) and (4), we will treat Case mismatches involving topicalized free relatives, cf. (3) in the text, and Case mismatches involving topicalized DPs, cf. (4), in a unified way. Both cases illustrate Case mismatches in CLLD chains.

- (5) (a) \*Agapo opjos m'agapai  
 love-1sg whoever-nom me-acc-loves
- (b) Agapo opjon m'agapai  
 love-1sg whoever-acc me-acc-loves  
 'I love whoever loves me'
- (6) (a) \*I kathigites agapane oli i fitites  
 the professors-nom love all-nom the-nom students-nom
- (b) I kathigites agapane olus tus fitites  
 the professors-nom love all-acc the-acc students-acc  
 'The professors love all the students'

(b) The Case possibilities on the relative pronoun heading CLLD free relatives depend on the Case of the main clause, as well as on the Case internal to the free relative. The examples in (7) contrast with those in (3) in that the Case internal to the free relative must appear in the relative pronoun if it is Genitive (cf. (7a–b)), as opposed to Nominative (cf. (3a–b)):

- (7) (a) Opju dosume to vradio, tha jini diasimos  
 whoever-gen give-perf-1pl the prize, fut become-perf-3sg famous  
 'Whoever we give the prize will become famous'
- (b) \*Opjos dosume to vradio, tha jini diasimos  
 whoever-nom give-perf-1pl the prize, fut become-perf-3sg famous

The examples in (8) show the particular status of Accusative with respect to Nominative and Genitive.<sup>4</sup> If Accusative is the Case checked by the free relative verb, it may appear on the relative pronoun if the matrix Case is Nominative, as shown in (8a):

- (8) (a) Opjos/Opjon den simbatho, den tha perasi to mathima  
 whoever-nom/acc not like-1sg, not fut pass the course  
 'Whoever I don't like will not pass the the course'

<sup>4</sup> As an anonymous reviewer has pointed out, the genitive *opjou* in (7a) can be replaced by the prepositional phrase *s'opjon* 'to + who-acc'. Such prepositional phrases can normally appear instead of adverbial genitives in Modern Greek. In this paper we will not explore the behaviour of these prepositional phrases in CLLD chains. Let us note though that the PP counterpart of (7a) in (i):

- (i) ?S'opjon dosume to vradio, tha jini diasimos  
 to-who-acc give-perf-1pl the prize, fut become-perf-3sg famous  
 'Whoever we give the prize will become famous'

is worse, though grammatical.

- (b) \*Opju epileksume, i Anna tha tu dosi ena vradio  
 whoever-gen choose-perf-1pl, the Anna fut him(cl)-gen- give-perf-3sg a prize
- (c) Opjon epileksume, i Anna tha tu dosi ena vradio  
 whoever-acc choose-perf-1pl, the Anna fut him(cl)-gen- give-perf-3sg a prize

On the contrary, as shown by (8b–c), Accusative must surface on the relative pronoun if the matrix Case is Genitive. Our central claim concerning the Greek free relative paradigm presented in (3) and (7)–(8) is that it cannot be accounted for in the strict limits of syntax, and that the morphological specification of Case in the relative pronouns, in interaction with a number of syntactic processes, is finally responsible for the facts in (3), (7)–(8). A syntactic account of the fact that the relative pronoun must surface in Genitive Case if the verb in the free relative has a Dative Case feature (cf. the contrast in (7)) can be provided on the basis of the distinction between **inherent/structural** Case.<sup>5</sup> Assuming that Dative is an inherent Case, tied to  $\Theta$ -marking (cf. Chomsky 1986b), and involves a +Interpretable feature, we can make sense of why the Genitive form of the relative pronoun, i.e., the expression of Dative abstract Case, is obligatory in free relatives involving a *wh*-chain with Dative Case. The deletion of the relevant Case would lead to a violation of Full Interpretation.<sup>6</sup> The behaviour of the Accusative, however, does not allow for a treatment in terms of the inherent/structural distinction. Whatever choice we make for Accusative, i.e., be it either inherent or structural, we would not expect the possibility of the relative pronoun surfacing in Accusative to depend on whether the CLLD matrix chain involves Nominative, cf. (8a), or Genitive, cf. (8b).

<sup>5</sup> In Greek, the indirect object shows the same Case as adnominal Genitives, as shown in (i):

- (i) a. Edosa tis Marias to vivlio  
 gave-1s the-gen Maria-gen the-acc book-acc  
 'I gave Maria the book'
- b. To vivlio tis Marias  
 the book the-gen Maria-gen  
 'Maria's book'

We refer to the Case occurring in (ia–b) as Genitive, although, as the example in (ib) indicates, this Case can also check Dative Case.

<sup>6</sup> A detailed presentation of this account falls beyond of the scope of this paper. A number of questions remain unanswered at this point. For instance, there is no conclusive evidence for the inherent status of Dative in Greek. In addition, the ban on deletion of forms expressing inherent Dative in free relatives based on the +Interpretable status of this Case raises the question of why some form of recoverability under deletion is not at play in these cases.

## 2. The analysis

In this paper, we claim that Case mismatches are mainly due to the fact that morphology, where word-formation occurs, is independent from syntax. We follow the approach according to which words are built within morphology and enter syntax already marked for their morphological features, Case being one of these features (cf., among others, Chomsky 1995).

In our analysis, we take advantage of the following assumptions put forward by Ralli (1997; to appear):

(a) in a modularly built grammar, morphology is an independent module interacting with syntax on several aspects within the computational system of the language faculty, cf. Di Sciullo (1996),

(b) general linguistic properties, inherent to human language, may be represented as features. These features belong to a feature theory module interacting with the grammatical modules, that is morphology, syntax and phonology, and

(c) inflectional features constitute the morphological expression of some of these features, but not all of them are visible to syntactic operations.

According to these assumptions, Case is an abstract universal notion belonging to a feature theory module. The encoding of Case takes place within an autonomous morphological module; syntax manipulates only the Case information that is syntactically relevant.<sup>7</sup> This means that, in languages, there could be instances where a word can be morphologically marked for a particular Case value without any need for syntax to provide licensing (e.g., checking) for this value. The assumptions above also predict that all members of the module of features are not overtly realized, i.e., morphologically expressed, in every single language and each grammatical component chooses the features that are appropriate for its own purposes. Since it depends on the particular language to choose the features for its own morphological system, the features of Case, Number and Gender, which are overtly realized in Greek inflection, belong to the morphological make-up of Greek words, but some of these features could eventually be absent from the morphological system of another

<sup>7</sup> That Case is an abstract universal notion from which one can deduce concrete forms in morphology, as well as more or less concrete notions in the other grammatical modules, was firstly proposed by Hjelmslev (1935, 85). He also states that these notions cannot be determined in isolation, but only throughout a system of syntactically and semantically determined variants.

language, e.g., Mohawk. It is further predicted that the three features do not have to be equally visible to syntax.<sup>8</sup>

Let us proceed by examining firstly the notion of non-movement chain formation, and the non-distinctiveness condition on the features involved in this sort of chains proposed in (liv).

## 2.1. Coindexation and feature interpretability

In this section, we would like to propose that the Chain Condition in (1) must involve constituents that are nondistinct in +Interpretable features. Thus, our claim is that the nondistinctiveness of the features involved in CLLD-chains should be relativized according to the +/–Interpretable distinction in Chomsky (1995). On the basis of this distinction, Case is syntactically characterized as –Interpretable, while Gender and Number are +Interpretable.<sup>9</sup> A –Interpretable feature, i.e., Case, is invisible to the C-I level, as opposed to the +Interpretable features of Gender and Number which belong to the features visible to this level. This distinction between +Interpretable and –Interpretable features allows us to derive the contrast in grammaticality be-

<sup>8</sup> As may be expected by such assumptions, the overt manifestation of Cases is not isomorphic with the set of functions that these Cases may express. For instance, more than one functions are expressed by Genitive case in Greek, e.g., indirect object, adnominal relation, cf. footnote 5.

<sup>9</sup> Notice that the +/–Interpretable opposition could also distinguish different types of cases. For instance, structural cases, such as Accusative or Nominative, are, in the general case, –Interpretable, whereas inherent cases could be considered +Interpretable, since they are, by definition, closely related to  $\Theta$ -marking. However, even the most structural cases may sometimes be related to a semantic interpretation, and thus behave as +Interpretable features, as the following Latin example seems to indicate, cf. Blake (1994, 32):

- (i) Vado Romam  
 go-1sg Rome-acc (destination)  
 ‘I go to Rome’

The occasional behaviour of Accusative as an inherent Case can be explained by assuming that a constituent may be associated to both a structural and an inherent Case in certain contexts, as argued in Torrego (1998) (but see Babby (1980), Yip – Maling – Jackendoff (1987), and McCreight (1988) for the opposite view). In section 2.3.1, we propose that Case, both in its structural and inherent manifestations, is never a +Interpretable feature. The fact that inherent Case may show certain properties of +Interpretable features can be derived from its close relation to  $\Theta$ -marking.

tween the sentences in (9) involving Gender and Number mismatches, and (3)–(4) involving Case mismatches:

- (9) (a) \*Opja                    m'agapai,            ton                    agapo  
           whoever-fem-nom    me-acc-loves,    him-masc-acc    love-1sg  
           'I love whoever loves me' (or *Whoever loves me, I love him*)
- (b) \*O    fititis,                    i    kathijites            tus                    agapane olus  
           the student-nom-sg, the professors-nom    them-acc-pl love-3pl all-acc-pl  
           'The professors love all the students'  
           (or *The students, the professors love them all*)

Under the assumption that Gender and Number, as opposed to Case, are visible to the C-I interface, and thus +Interpretable features, we can rule out the sentences in (9) based on the reformulation of Baker's fourth condition on non-movement chain formation, cf. (1), in (10).

- (10) X and Y are non-distinct in +Interpretable features (i.e., Number, Gender, Person, etc.)

(10) excludes –Interpretable features, such as Case, from the matching conditions on the non-movement chain formation, allowing for the grammaticality of sentences such as (3)–(4). The exclusion of –Interpretable features from whatever is the exact mechanism relating a CLLD-constituent to its clitic makes sense if we view this mechanism as a reflex of a C-I relation, which, by definition, can only be expressed in terms of features interpretable at that interface. Notice that Case as well as Number and Gender are involved in the word-formation procedure of both the dislocated constituent and the clitic. However, they are grouped differently in morphology. For instance, Case and Number belong to the same inflectional cluster representing the nominal ending, while Gender characterizes the stem.<sup>10</sup> On the other hand, in clitics and determiners, the same portmanteau morpheme hosts all three features. This difference in use of the same features by morphology and syntax may be viewed as further proof that the two modules are independent and that each module provides its own means to manipulate the features appropriate for its purposes.

Having said this, the following questions still need to be answered:

<sup>10</sup> See Ralli (1994) for more details on this claim and for an analysis of the Gender feature as an inherent marker of Greek stems.

(a) Why is nominative the only alternative Case value for CLLD constituents, as shown in (11)–(12), which involve a dative and an accusative dislocated constituent respectively?<sup>11</sup>

- (11) (a) (?)O            Pavlos,        tu            pirane            to        pedi  
                           the-nom Paul-nom,    him-gen    took-away-3pl    the    kid
- (b) Tu            Pavlu,        tu            pirane            to        pedi  
                           the-gen Paul-gen,    him-gen    took-away-3pl    the    kid  
                           ‘They took the kid away from Paul’
- (c) \*Ton        Pavlo,        tu            pirane            to        pedi  
                           the-acc Paul-acc,    him-gen    took-away-3pl    the    kid
- (12) (a) (?)I        fitites,            i        kathijites            tus            agapane    olus  
                           the students-nom,    the professors-nom    them-acc    love-3pl    all-acc  
                           ‘The professors love all the students’  
                           (or *The students, the professors love them all*)
- (b) Tus        fitites,            i        kathijites            tus            agapane    olus  
                           the students-acc,    the professors-nom    them-acc    love-3pl    all-acc  
                           ‘The professors love all the students’  
                           (or *The students, the professors love them all*)
- (c) \*Ton        fititon,            i        kathijites            tus            agapane    olus  
                           the students-gen,    the professors-nom    them-acc    love-3pl    all-acc

(b) Why do Nominative, Accusative, and Genitive behave differently in free relatives contexts, which combine a CLLD-chain and a *wh*-chain, cf. section 1? We deal with the questions in (a) and (b) in sections 2.2 and 2.3, respectively.

<sup>11</sup> As shown by the following example, the sentence becomes ungrammatical if the dislocated constituent appears in Vocative.

- (i) \*Pavle,        tu            pirane            to        pedi  
                           Paul-voc,    him-gen    took-away-3pl    the    kid

Vocative is the fourth value of the Greek Case system, but it will not be considered here because of its special character: it is the case form of “address” and marks constituents that stand outside construction, bearing no relation of dependents to heads. On the other hand, as shown by Kurylowicz (1964, 188), Nominative, Accusative and Genitive are the “most” grammatical cases, that is the cases closely related to grammatical relations.



## 2.2. Nominative as a default Case value

With respect to the first question above, we would like to claim that Nominative acts as a default Case value in CLLD contexts. Since non-movement chain formation does not involve identity of –Interpretable features, as claimed in (10), any Case value could be possible in CLLD contexts. Furthermore, CLLD contexts do not involve Case-checking configurations for the dislocated constituent, and we might expect checking of any Case feature on a CLLD constituent to be excluded, as opposed to what we have seen in Greek. The data we have considered so far show that the dislocated constituent may have either the same Case value as the one on the clitic, or Nominative (cf. (11)–(12)). Our proposal then is that Nominative is assigned within morphology and its presence in syntax can be considered as that of a default Case value which becomes possible for two reasons:

- (a) dislocated constituents and their doubling clitics do not need to agree in Case (cf. (10)), and
- (b) all Greek nominals must bear a Case value, as a consequence of their morphological make-up.

The fact in (b) leads to the possibility of the presence of a default value in nominals for which no particular value is syntactically justified. There are many pieces of independent evidence which square quite well with the idea that Nominative should be considered as the default option in the Greek Case system. For instance, Nominative has always been the Case outside construction in Greek, the Case for “naming” (*onomazo* in Ancient Greek). According to Humbert (1960, 249), it can be conceived independently of any grammatical/syntactic relations. On the other hand, Jakobson (1958) attributes the close relationship between Nominative and the concept of topichood (an observation that goes back to Aristotle) to the fact that Nominative is the Case value with less relational content.<sup>12</sup> Thus, it is not unreasonable to claim that in a non-checking syntactic configuration, such as the one involving left-

<sup>12</sup> The idea that the default use of Nominative is related to the notion of topichood is also exploited by Alexiadou – Varlokosta (1996). In their paper, the authors characterize left-dislocated constituents as instances of **hanging topics** marked for a default Nominative value (1996, 20). They claim that non-matching left-dislocated free relatives are also instances of hanging topics. According to Alexiadou – Varlokosta (1996, 20) “the default Nominative case and the presence of a resumptive clitic (cf. (i)) are characteristics of hanging topics”:

- |     |             |         |              |            |
|-----|-------------|---------|--------------|------------|
| (i) | Opjos       | argisi  | ton          | timorun    |
|     | whoever-nom | is late | him-(cl)-acc | punish-3pl |

dislocated constituents, Nominative could appear as the only alternative value expressed by these constituents, beside Genitive (11b) or Accusative (12b), depending on the case of the doubling clitic. The proposal that CLLD contexts constitute a kind of non-checking syntactic configuration, where Nominative may be used as a default option is also supported by the examples below: it is shown that when there is a clear-cut checking requirement for the presence of another case value, i.e., Genitive or Accusative: Nominative is not possible.

- (13) (a) Opjon tu dosis to onoma mu, tha ton voihiso  
 whoever-acc him-gen give-perf-2sg the name my-gen, fut him-acc help-perf-1s  
 'I'll help whoever you give my name to'
- (b) Opju tu dosis to onoma mu, tha ton voithiso  
 whoever-gen him-gen ..... him-acc .....
- (c) Opjos tu dosis to onoma mu, tha ton voithiso  
 whoever-nom him-gen ..... him-acc .....
- (d) \*Opjos tu dosis to onoma mu, tha voithiso  
 whoever-nom him-gen ..... fut help-perf-1sg
- (e) \*Opjos dosis to onoma mu, tha voithiso  
 whoever-nom ..... fut help-perf-1sg

As shown in (13a), Greek free relative clauses can also contain clitics agreeing with the relativized constituent (cf. Horrocks–Stavrou 1987). With two clitics, one in the free relative and another in the main clause, we have three possible Cases for the relative pronoun: Accusative (13a), Genitive (13b), and also Nominative (13c), although somehow marginally. However, Nominative for the relative pronoun is completely excluded if one of the clitics is missing, as shown by the ungrammaticality of (13d–e). This shows that Nominative is only possible on the relative pronoun when the pronoun is coindexed with clitics, i.e., *tu* and *to* in (13c). When one of the clitics is missing, which is tantamount to saying that there is a movement chain which requires a Case corresponding to the one checked against the verb without a clitic, then Nominative is not an option any more. Thus, we can conclude that Nominative is a sort of default option, only available when the relative pronoun does not check any Case, either in the free relative or in the main clause. Notice, however, that this default use of Nominative for syntactic reasons does not coincide with what can be a morphologically unmarked option, since Nominative in Greek very

---

However, non-matching free relatives in left-dislocated position do not necessarily show Nominative case, as can be seen from examples like (7a) or (8c)). This indicates that the characterisation of free relatives as hanging topics is incorrect.

often constitutes the marked form of the nominal paradigm. Consider, for instance, the very common masculine nouns in *-is* or *-as* (cf. (14)), where the Nominative singular is expressed by the affix *-s* whereas the Genitive is not overtly expressed (or expressed by the  $\emptyset$ -affix). This is another example in favour of the independence of the morphological module:

- |      |            |          |           |
|------|------------|----------|-----------|
| (14) | Nominative | Genitive |           |
|      | fititi-s   | fititi   | 'student' |
|      | tamia-s    | tamia    | 'cashier' |

### 2.3. Greek free relatives

Case resolution in Greek CLLD-free relatives constitutes an interesting case study from a typological point of view. The reason is that Greek is a language which combines overt case marking with CLLD constructions. Most of the literature on Case related effects in free relatives deals either with languages that have clitics and CLLD constructions, but not overt case marking like Spanish or Catalan (e.g., Hirschbüler–Rivero 1981; Suñer 1984) or with languages that have rich case systems but not clitics (e.g., Groos–van Riemsdijk 1981; Harbert 1983; McCreight 1988). From the consideration of Greek free relatives in CLLD constituents, we can draw further evidence for our claim that non-movement chains may involve different values for –Interpretable features in different members of the chain. Consider, for instance, the behaviour of Genitive case as illustrated in (15):

- (15) (a) Opjos ftasi protos, tha tu dosume ena vradio  
 whoever-nom arrives-perf-3sg first, fut him(cl)-gen give-perf-1pl a prize  
 'Whoever arrives first, we'll give him a prize'
- (b) Opjon epileksume, i Anna tha tu dosi ena vradio  
 whoever-acc choose-perf-1pl, the Anna fut him(cl)-gen give-perf-3sg a prize  
 'Whoever we choose, Anna will give him a prize' (cf. (8c))
- (c) \*Opjos dosume to vradio, tha jini diasimos  
 whoever-nom give-perf-1pl the prize, fut become-perf-3sg famous  
 'Whoever we give the prize will become famous' (cf. (7b))
- (d) \*Opjon dosume to vradio, i Anna tha ton proslavi  
 whoever-acc give-perf-1pl the prize, the Anna fut him(cl)-acc hire  
 'Whoever we give the prize, Anna will hire him'

(15a–b) show that the Case checked inside the free relative may surface on the relative pronoun, in spite of the fact that the free relative forms a non-

movement chain with a genitive clitic.<sup>13</sup> On the other hand, (15c–d) show that the matrix Case, Nominative in (15c) and Accusative in (15d), cannot surface on the relative pronoun when there is a dative case checked internally to the free relative. This state of affairs straightforwardly follows from the reformulation of Baker's theory proposed in (10). In non-movement chains, as opposed to *wh*-chains, there is no requirement on identity of –Interpretable features, allowing for Case mismatches in (15a–b), but not in (15c–d). In addition, the contrast between obligatory Case matching in argumental free relatives observed in (5), as opposed to its optionality in left-dislocated position also follows partially from (10). CLLD constructions, such as the one in (16b), involve a clitic which checks the Case of the matrix verb, and, according to (10) this Case does not need to surface on the relative pronoun:

- (16) (a) \*Agapo opjos m'agapai  
 love-1g whoever-nom me-acc-loves
- (b) Opjos m'agapai, ton agapo  
 whoever-nom me-acc-loves, him-(cl)-acc love-1sg  
 'Whoever loves me, I love him'

On the other hand, when the free relative is in argument position, the matrix Case must be checked by the free relative constituent, so that the Case internal to the free relative cannot surface on the relative pronoun. The fact that non-matching free relatives involve in the general case a pronominal element which checks the matrix Case seems to be quite general. Consider the following example from Polish involving Accusative matrix case and Dative in the free relative, cf. McCreight (1988, 94):<sup>14</sup>

<sup>13</sup> In certain cases (cf. (22) below), genitive case may alternatively surface on the relative pronoun. Thus, (i) is also possible:

- (i) Opjou ftasi protos, tha tu dosume ena vradio  
 whoever-gen arrives-perf-3sg first, fut him(cl)-gen give-perf-1pl a prize  
 'Whoever arrives first, we'll give him a prize'

<sup>14</sup> We abstract away from the fact that in many languages the pronominal element checking matrix Case may be absent if there is a syncretic form of the relative pronoun that can accommodate matrix and free relative Cases, as in the Polish example in (i):

- (i) kupiłam co było w sklepie  
 bought what-nom/acc was in the store  
 'I bought what was in the store'

For space reasons, we will not discuss either cases of hierarchical resolution, cf. Harbert (1983), and McCreight (1988). Our discussion in section 2.3.1, however, is undoubtedly related to this effect.

- (17) Marysia postanowiła kupić to, czemu Janek się przyglądał  
 Maria decided to-buy it-acc, that-dat Janek self stared-at  
 'Maria decided to buy what Janek was staring at'

Cheila-Markopoulou (1991) has also argued in detail that free relatives in Medieval Greek are strictly non-matching. However, her examples always involve a matrix clitic checking the matrix Case, as for instance, the one in (18):

- (18) etimos (...) na ton polemiso opjos na ipi oti  
 ready na him-(cl)-acc fight-perf-1sg whoever-nom na says-perf-3sg that  
 esfala  
 be-wrong-perflsg  
 'ready to fight whoever says that I was wrong'

Perhaps the most interesting fact about Greek dislocated free relatives is that they optionally allow for matrix Case on the relative pronoun, as shown in (19) (cf. (3)):

- (19) (a) Opjos m'agapai, ton agapo  
 whoever-nom me-acc-loves, him-(cl)-acc love-1sg  
 (b) Opjon m'agapai, ton agapo  
 whoever-nom me-acc-loves, him-(cl)-acc love-1sg  
 'Whoever loves me, I love him'

We would like to claim that (19b) should be viewed as a manifestation of the poorly understood phenomenon of case attraction (cf. Groos-van Riemsdijk 1981; Harbert 1983). Groos-van Riemsdijk (1981) define **case attraction** as "the term referring to situations in which the *wh*-phrase agrees in case with its antecedent, or—in the case of free relatives—receives its case marking from the matrix clause," as shown in the Classical Greek example in (20):

- (20) aksioi tes eleutherias, hes (for hen) kektesthe  
 gen gen acc  
 'worthy of the freedom which you possess'

(X.A. 1.7.3, cf. Smyth 1963, 567)

Thus, the genitive case of *opjon* in (19b) is assimilated as a case of attraction by the matrix clitic *ton*. CLLD free relatives in Greek share two crucial properties of case attraction contexts: (a) case attraction is always optional, as shown by (19) and (20), and (b) case attraction seems to obey a case hierarchy. As argued in Harbert (1983), case attraction in Classical Greek follows the hierarchy in (21):

(21) acc < dat < gen<sup>15</sup>

The facts presented in (7)–(8) also suggest at first sight some form of Case hierarchy affecting the Case possibilities of relative pronouns in CLLD-free relatives. However, a more detailed look reveals that the relation between Cases in Greek CLLD-free relatives cannot be described in terms of Case hierarchy. The table in (22) summarizes the resolution of all the possible combinations of Cases in the relevant contexts:<sup>16</sup>

|      |     |         |         |         |
|------|-----|---------|---------|---------|
| (22) | +cl | nom     | acc     | gen     |
|      | –cl |         |         |         |
|      | nom | nom     | nom/acc | nom/gen |
|      | acc | nom/acc | acc     | acc     |
|      | gen | gen     | gen     | gen     |

The case in the +cl row corresponds to the case in the non-movement chain, whereas the case in the –cl column corresponds to the case in the *wh*-chain. According to (22), Nominative may attract into other cases (cf. footnote 17), Accusative only attracts into Nominative, and Genitive never attracts into any other case. This situation cannot be expressed in terms of a hierarchical relation among Cases, because Accusative case attracts into Nominative, which in its turn attracts into Genitive, but it does not attract into Genitive which in its turn does not attract into Nominative. In the next section, we will introduce a morphological analysis of the Greek Case system which, in interaction with the syntactic properties of Case, accounts for the situation summarized in (22). This analysis provides evidence for viewing morphology as a component of grammar which may involve relations among feature values in a way quite different from those attested in syntax.

### 2.3.1. Morphological make-up of Greek relative pronouns

In section 2.1, we have seen that the use of a default nominative option in CLLD contexts leads to the conclusion that syntax may use actively only partial morphological information in its operations. Our analysis crucially shows

<sup>15</sup> Nominatives behave in a quite strange way with respect to the hierarchy in (21). According to Harbert (1983, 62), masculine and feminine nominatives do not attract into other cases, but neuter nominatives/accusatives do attract. This fact suggests that the hierarchy in (21) should be, at least partially, morphologically motivated.

<sup>16</sup> Cf. Androutsopoulou–Español–Echevarría (1995).

that, in specific contexts, syntax ignores pieces of information coming from fully inflected items, i.e., Case, while it manipulates some other information characterizing the same items, i.e., Gender, and Number. As we will see below, the same applies to the case of free relatives: the morphological Case marking may be richer than one would expect based on syntactic considerations. This is tantamount to saying that syntax may only partially use the morphological information on Case found in the lexical items introduced in the syntactic derivation.

We would like to focus on two aspects of the table in (22) which suggest that Nominative and Accusative should be grouped together in opposition to Genitive, and that this opposition is not only syntactic in nature.<sup>17</sup> First, Genitive case cannot be attracted into Nominative nor Accusative. This fact may be derived from the syntactic characterisation of Dative (morphologically Genitive) as inherent Case. If morphological Genitive is the surfacing form in

<sup>17</sup> In this paper, we will not discuss in detail why Nominatives may always be attracted into other Cases, cf. (22). We think that this is related to the fact that preverbal subjects are topics in Modern Greek (cf. Philippaki-Warbuton (1985), among others). Topic subjects may form a non-movement chain with a *pro* empty category in subject position, as shown in (i), cf. (19b):

- (i) [[Opjoni<sub>i</sub>    [pro<sub>i</sub>    m'agapai]]    ton<sub>i</sub>; agapo]  
       whoever                    me-loves            him-(cl)-acc-love-1sg

The chain *opjon-pro* in (i) behaves in the relevant respects as a CLLD chain allowing for attraction into other Case, e.g., Accusative in (i). This analysis is also compatible with the facts concerning subject free relatives. Thus, we assume that (iia) has the structure in (iib):

- (ii) a. Opju            dosume            to vravio, tha jini            diasimos  
       whoever-dat give-perf-1pl the prize, fut become-3sg famous  
       'Whoever we give the prize, he will become famous'  
       b. [Opju dosume to vravio]<sub>i</sub>; pro<sub>i</sub> tha jini diasimos

The fact that the Case checked inside the free relative may always surface on the relative pronoun in subject free relatives follows from the configuration in (iib), which does not involve checking of the matrix Nominative case by the free relative. The particular status of Nominatives in this respect is not exclusive to Greek, and according to our analysis should be attested in other pro-drop languages. McCreight (1988, 94) notes that in Polish only the Nominative checked in the matrix clause may be suppressed when the relative clause is the subject of the matrix clause, as shown in the examples in (iii):

- (iii) a. To czego Janek oczekiwał w końcu zdarzyło się  
       'That which Janek expected finally happened'  
       b. ?Czego Janek oczekiwał w końcu zdarzyło się  
       'What Janek expected finally happened'

Greek of an inherent Dative Case (cf. footnote 5) checked by the verb in the free relative, attraction by the matrix Case, as in the ungrammatical example in (23), cf. (15c-d), would entail in minimalist terms that the inherent Case feature checked in the free relative is erased, which is presumably incompatible with the inherent nature of the Dative case.<sup>18</sup>

- (23) (a) \*Opjos dosume to vravio, tha jini diasimos  
 whoever-nom give-perf-1pl the prize, fut become-perf-3sg famous  
 'Whoever we give the prize will become famous' (cf. (7b))
- (b) \*Opjon dosume to vravio, i Anna tha ton proslavi  
 whoever-acc give-perf-1pl the prize, the Anna fut him(cl)-acc-hire  
 'Whoever we give the prize, Anna will hire him'

However, there is no strong evidence supporting the inherent status of Dative case (morphologically genitive) in Modern Greek. Furthermore, there seems to be evidence for the contrary. If the Genitive on *tu orfanotrofiu* 'the orphanage' was the morphological expression of an inherent Dative in (24), we would expect this Case marking to be preserved in the nominalized counterpart. The ungrammaticality of (24b) shows that this expectation is not fulfilled.<sup>19</sup>

- (24) (a) Dosame tu orfanotrofiu polla ruha  
 gave-perf-1pl the-gen orphanage-gen many clothes
- (b) \*I dorea tu orfanotrofiu (under a non-possessive interpretation)  
 the gift the-gen orphanage-gen

<sup>18</sup> The assumption that inherent Case cannot be erased is supported by the typical behavior of inherently Case marked constituents under passivization, as in the examples in (i) showing that dative case in German is retained under passivization:

- (i) a. \*er wird geholfen  
 he-nom is helped  
 b. ihm wird geholfen  
 he-dat is helped cf. Belletti (1988)

The fact that inherent case features cannot be erased from syntactic derivations before Spell-Out does not necessarily imply that they are +Interpretable features. We would like to claim that it is the  $\Theta$ -relation the inherent Case is associated with, and not the Case feature per se, which is +Interpretable.

<sup>19</sup> It may be the case that the licensing of Dative arguments involves both inherent and structural cases, along the lines suggested in Torrego (1998). This would account for the ungrammaticality of the nominalized counterpart where the hypothetical structural case could not be checked. We will not consider this possibility in this paper.



- (c) I dorea sto orfanotrofio  
 the gift to-the-acc orphanage-acc

Second, as noted in section 1.1, cf. also (22), Accusative may attract into Nominative but not into Genitive, as shown in (8), repeated here as (25):

- (25) (a) Opjos den simbatho, den tha perasi to mathima  
 whoever-nom not like-1sg, not fut pass the course  
 'Whoever I don't like will not pass the course'
- (b) \*Opju epileksume, i Anna tha tu dosi ena vradio  
 whoever-gen choose-perf-1pl, the Anna fut him(cl)-gen give-perf-3sg a prize
- (c) Opjon epileksume, i Anna tha tu dosi ena vradio  
 whoever-acc choose-perf-1pl, the Anna fut him(cl)-gen give-perf-3sg a prize

The contrast in (25) is obviously not reducible to a simple characterisation of Accusative as either inherent or structural Case, since its PF realization depends on the matrix Case. We would like to propose that a special morphological feature [+/-Genitive] that crosscuts the Greek morphological system of Case values interferes with the syntactic operation of checking as an instance of morphology-syntax interaction. The syntactic visibility of this morphological feature makes it possible for a constituent bearing an overt Accusative, or an overt Nominative, i.e., a constituent marked as [-Genitive], to be checked in a syntactic context requiring one of the two values (cf. (25a)). As opposed to this, in a syntactic context requiring [+Genitive], only constituents morphologically marked for Genitive can be checked (cf. (23)).

The ungrammaticality of (25b) follows from the fact that the [+Genitive] form *opjou* cannot check accusative case on *epileksume*. On the contrary, the Nominative form *opjos* in (25a) can check accusative case on the free relative verb. Furthermore, the character of overt Genitives as resisting case attraction directly follows from the present proposal, even in the case in which Dative (morphologically expressed as Genitive) turns out to be structural Case in Modern Greek.

The idea of grouping together Nominative and Accusative, as opposed to Genitive, is supported from several considerations. Firstly, as noted by Humbert (1960, 248), the history of Greek has always shown a direct relation between Nominative and Accusative. On the basis of head-dependent relations, Genitive is considered to be an adnominal Case as opposed to the adverbial

Cases of Nominative and Accusative.<sup>20</sup> On the other hand, in the evolution of the language, there are many instances of nouns which have adapted their irregular nominative form analogically to the 'more regular' accusative form:

|      |                    |                    |                     |
|------|--------------------|--------------------|---------------------|
| (26) | Fifth century B.C. | Third century A.D. | (cf. Browning 1969) |
|      | Nominative         | Accusative         | Nominative          |
| (a)  | pater<br>'father'  | patera             | → pateras           |
| (b)  | meter<br>'mother'  | metera             | → metera            |

Case syncretism in Greek also functions on the basis of this opposition. Neuter nouns show the same inflected form in both Nominative and Accusative while their Genitive form is different. Consider, for example, the inflectional paradigm of the neuter noun *soma* 'body' below:

|      |     |          |         |
|------|-----|----------|---------|
| (27) |     | Singular | Plural  |
|      | Nom | soma     | somata  |
|      | Acc | soma     | somata  |
|      | Gen | somatos  | somaton |

Moreover, some derived words today, e.g., the diminutives in *-aki*, do not have an inflected form for Genitive in Singular and Plural, while they are fully inflected for Nominative and Accusative:

|      |     |               |   |       |   |                  |
|------|-----|---------------|---|-------|---|------------------|
| (28) | (a) | pedaki        | < | ped-  | + | -aki             |
|      |     | 'small child' |   | child |   | small-nom/acc-sg |
|      | (b) | pedakia       | < | ped-  | + | -akia            |
|      |     |               |   | child |   | small-nom/acc-pl |
|      | (c) | *pedakiu      | < | ped-  | + | -aki             |
|      |     |               |   | child |   | small-gen-sg     |
|      | (d) | *pedakion     | < | ped-  | + | -akion           |
|      |     |               |   | child |   | small-gen-pl     |

The idea of Case checking being influenced by the language-dependent morphological feature of [+/-Genitive] allows us to explain the situation presented in (22). The question is, however, whether a word bearing a morphologically overt Accusative, or an overt Nominative, can freely check one or the other

<sup>20</sup> On syntactic grounds, the same division of cases is also proposed by Simon of Dacia for the Case system of Latin. Simon of Dacia divides the cases into those which express a substance to substance relation and those which do not:

|                        |      |      |      |                      |
|------------------------|------|------|------|----------------------|
|                        | Nom. | Acc. | Gen. |                      |
| Substance to substance | -    | -    | +    | (cf. Blake 1994, 37) |

value. In other words, can Nominative forms alternate with Accusative forms in all appropriate contexts? As (29) shows, this is not possible:

- (29) (a) \*Ton        Jani        efere        luludia  
           the-acc    John-acc    brought    flowers-acc-pl
- (b) O            Janis        efere        luludia  
           the-nom    John-nom    brought    flowers-acc-pl  
           ‘John brought flowers’

In (29a), the DP *ton Jani*, is morphologically marked for Accusative, but the sentence is ungrammatical because only Nominative is the syntactically checked case for subjects. At this point, we would like to claim that there is a M(orphological) M(odule) Constraint according to which morphological forms should be as transparent as possible with respect to syntactic operations.

- (30) \* Opacity on syntactic operations

The MM constraint in (30) captures the observation that unexpected Case values only arise whenever two syntactic relations mediating Case require different Case values on a concrete lexical item. It also introduces the notion that Case-marked morphological representations are “optimal” expressions of syntactic Case-checking requirements. This allows for an account of why both Nominative and Accusative Case values are possible in (31), but not in (29).

- (31) Opjos/opjon        den simbatho, den tha        perasi        tis eksetasis  
           whoever-nom/acc not like-1sg    not fut-pass-3sg the exams  
           ‘Whoever I don’t like, won’t pass the exams’

In (29), Accusative case marking on the subject violates (30), and, thus, (29a) is excluded. In (31), however, any of the choices, i.e. Nominative or Accusative, leads to a violation of (30): Accusative case on the relative pronoun induces opacity with respect to the chain formed by the dislocated free relative and the *pro* empty category in subject position (cf. footnote 17), whereas Nominative case induces opacity on the Case checked by the relative pronoun inside the free relative, i.e. Accusative. Therefore both choices equally violate (30), and this is why both cases, Nominative and Accusative, are equally acceptable in this context.<sup>21</sup>

<sup>21</sup> In the general case, as observed in (5), the relative pronoun in free relatives in argumental position cannot show the Case internal to the free relative, as shown in (i), cf. (5):

### 3. Morphological syncretism and syntactic syncretism

In this section we would like to briefly discuss our proposal about the interaction between syntax and morphology in the general framework of Case resolution. We would like to note that our proposals about the way in which two cases are accommodated in dislocated Greek free relatives can be understood as a particular case of syncretic resolution of Case conflicts. Morphological syncretic resolution of Case conflicts is a general phenomenon in free relative contexts. Taraldsen (1981) shows for German that while neither Nominative *wer* 'who' nor Accusative *wen* 'who' can accommodate the matrix Nominative case and the free relative Accusative case in the examples in (32), the syncretic Nominative/Accusative form *was* 'what' can accommodate the two cases, as shown in (33):

(32) (a) \*Ich zerstöre wer mich ärgert

(b) \*Ich zerstöre wen mich ärgert  
'I destroy who annoys me'

(33) Ich zerstöre was mich ärgert  
'I destroy what annoys me'

The ability of *was* to accommodate Nominative and Accusative in free relative contexts is a direct consequence of the properties of its morphological Case marking. It can be argued on the basis of (32)–(33) that *was* is underspecified with respect to the Case feature discriminating Nominative from Accusative

---

(i) \*Agapo opjos m'agapai  
love-1sg whoever-nom me-acc-loves

The example in (i) crucially differs from the CLLD cases in that in (i), the relative pronoun is related to two checking configurations. In the CLLD cases, there is only one checking relation, that is, the one taking place inside the free relative. In this section, we have claimed that forms morphologically marked as [–Genitive] may check either Nominative or Accusative. The example in (i) shows that they cannot check both in the same derivation. This constraint follows in minimalist terms from the very nature of the checking relation. Once an item checks its Case feature, this feature cannot enter into other Case checking relations. The analysis of Case resolution in argumental free relatives falls beyond the scope of this paper. Nonetheless, we would like to note that the matching/non-matching character of Greek free relatives cannot be deduced from their syntactic position alone. The example in (ii) shows that if the free relative involves Dative case, matching is not possible:

(ii) \*Tha proslavo opjon dosume to vivlio  
fut hire-1sg whoever-acc give-perf-1pl the book  
'Whoever we give the book, I will hire him'

in German. This underspecification is at the root of the particular behaviour of *was* in (33). The proposal presented in this paper crucially makes use of the notion of syncretism, but in the Greek case at a syntactic level rather than at a morphological one. In Greek, syntax seems to use a part of the feature description provided by morphology, for instance, the [+/-Genitive] value argued for in the preceding section. Thus, morphologically non-syncretic forms such as Nominative *opjos* or Accusative *opjon* are rendered syncretic and suitable to accommodate different Case relations. This is in fact a kind of syncretic strategy for Case conflict resolution, which has a syntactic source.

#### 4. Conclusions

In this paper, we have shown that an attempt to link the Case forms directly to syntactic operations is unwieldy since grammatical relations need not be in a one-to-one correspondence with Case forms, e.g., Nominative forms in CLLD contexts. We argued that Case marking is handled within morphology and syntax may manipulate only partial Case information coming from morphology. Our analysis provides a strong argument for the existence of a morphological module, but also a confirmation for the interaction between morphology and syntax. The cases we considered here are instances in which either syntax does not use all the information provided by morphology, or the syntactic operations are affected by constraints coming from morphology.

#### References

- Alexiadou, A. – Varlokosta, S. 1996. The syntactic relevance of free relatives in Modern Greek. Ms. FAS Berlin and UPenn.
- Androutsopoulou, A. – Español-Echevarría, M. 1995. Case attraction in Classical Greek relative clause. Ms., UCLA.
- Babby, L. 1980. The syntax of surface case marking. In: Cornell Working Papers in Linguistics I. Cornell Linguistics Publications, Ithaca NY.
- Baker, M. 1996. The Polysynthesis Parameter. Oxford University Press, Oxford/New York.
- Belletti, A. 1988. The case of unaccusatives. In: Linguistic Inquiry 19: 1–34.
- Blake, B. 1994. Case. Cambridge University Press, Cambridge.
- Browning, R. 1969. Medieval and Modern Greek. Hutchinson, London.
- Catsimali, G. 1990. Case in Modern Greek: implications for clause structure. Ph.D. dissertation, University of Reading, Reading.

- Cheila-Markopoulou, D. 1991. Problems of diachronic syntax: free relatives in Medieval and Modern Greek. *Glossologia* 9-10 : 1-30.
- Chomsky, N. 1986. *Knowledge of language: its nature, origin, and use*. Praeger, New York.
- Chomsky, N. 1995. *The Minimalist Program*. The MIT Press, Cambridge MA.
- Cinque, G. 1990. *Types of A-bar dependencies*. MIT Press, Cambridge MA.
- DiSciullo, A. 1996. Atomicity and relatedness in configurational morphology. In: DiSciullo, A. (ed.): *Configurations essays on structure and interpretation*, 17-41. Cascadilla Press, Somerville.
- Groos, A. – van Riemsdijk, H. 1981. Matching effects in free relatives: a parameter of core grammar. In: Belletti – Brandi – Rizzi (eds): *Theory of markedness in generative grammar*, 171-216. Scuola Normale Superiore, Pisa.
- Harbert, W. 1983. On the nature of the matching parameter. In: *The Linguistic Review* 2 : 237-84.
- Hirschbuler, P. – Rivero, M. 1981. A unified analysis of matching and non-matching free relatives in Catalan. In: Burke, V. – Pustejovsky, J. (eds): *NELS 11*, 113-24. GLSA, Amherst MA.
- Hjelmslev, R. 1935. La catégorie des cas. Étude de grammaire générale I. In: *Acta Jutlandica* 7 : xij-184.
- Horrocks, G. – Stavrou, M. 1987. Bounding theory and Greek syntax: evidence for *wh*-movement in NP. In: *Journal of Linguistics* 23 : 79-108.
- Humbert, J. 1960. *Syntaxe Grecque*. Klincksieck, Paris.
- Jakobson, R. 1958/1971. Morphological inquiry for Slavic declension: structure of Russian case forms. Text in Russian, summary in English in 1971 version. In: Jakobson, R.: *Selected writings II: Words and language*, 154-83. Mouton, The Hague.
- Kuryłowicz, J. 1964. *The inflectional categories of Indo-European*. Carl Winter, Heidelberg.
- McCreight, K. 1988. *Multiple case assignments*. Ph.D. dissertation, MIT, Cambridge MA.
- Philippaki-Warbuton, I. 1985. Word order in Modern Greek. In: *Transactions of the Philological Society*, 113-43.
- Philippaki-Warbuton, I. 1990. The analysis of verbal complex in Modern Greek. In: *Studies in Greek Linguistics* 10 : 119-39.
- Philippaki-Warbuton, I. – Stavrou, M. 1986. Free relatives in Modern Greek. In: *Studies in Greek Linguistics* 7 : 117-34.
- Ralli, A. 1994. Feature representations and feature-passing operations in Greek nominal inflection. In: *Proceedings of the 8th symposium on English and Greek linguistics*.
- Ralli, A. 1997. Inflectional features and the morphological component hypothesis. Evidence from Modern Greek. Paper presented at Paper at 20th GLOW, Rabat.
- Ralli, A. to appear. On the morphological status of inflectional features: evidence from Modern Greek. In: Horrocks, G., Joseph, B. – Philippaki-Warbuton, I. (eds): *Themes of Greek linguistics II*. John Benjamins, Amsterdam.

- Smyth, H. 1963. Greek grammar. Harvard University Press, Cambridge MA.
- Suñer, M. 1984. Free relatives and the matching parameter. In: *The Linguistic Review* 3: 363–87.
- Taraldsen, T. 1981. Case-conflict in Norwegian topicalization. In: Burke, V. – Pustejovsky, J. (eds): *NELS 11*, 377–98. GLSA, Amherst MA.
- Torrego, E. 1998. The dependencies of objects. MIT Press, Cambridge MA.
- Tsimpli, I. 1990. The clause structure and word order in Greek. In: *UCL Working Papers in Linguistics* 2: 226–55.
- Tzartanos, A. 1946. *Neoelliniki Syntaxis tis Koinis Neoellinikis* [Greek syntax of Modern Greek Koine]. OESB, Athens.
- Yip, M., Maling, J. – Jackendoff, R. 1987. Case in tiers. In: *Language* 63: 217–50.

Address of the authors: Manuel Español-Echevarría  
Département de langues et linguistique  
Pavillion Charles-de Koninck  
Université Laval  
Québec, Québec G1K 7P4  
Canada  
Manuel.Espanol-Echevarria@lli.ulaval.ca

Angela Ralli  
Prof. of Linguistics  
Department of Philology (Linguistics Section)  
School of Humanities and Social Sciences  
University of Patras  
26500 Rio-Patras Greece  
aralli@atlas.uoa.gr





## ON THE INTERACTION BETWEEN MORPHOLOGY AND SEMANTICS: THE ITALIAN SUFFIX -ATA

LIVIO GAETA

### Abstract

Action nouns are often claimed to be sensitive to the actional properties of verbs. In this paper, an attempt will be made to consider the possible interactions between the morphological rules that form action nouns and the actional content of verbs. In this respect, a notion of internal and of external actionality of an affix will be distinguished, which are respectively responsible for the affix's selection properties and for its global semantics. The accurate analysis of the Italian suffix -ATA will reveal that both internal and external actionality play a crucial role in delimiting the input and in defining the semantics of the output.

### 1. Introduction

As has been pointed out by several authors (cf. Giacalone Ramat 1974; 1975; Ullmer-Ehrich 1977; Ehrich 1991; Bartsch 1981; 1986; ten Cate 1985; Brinton 1993), action nouns are sensitive to the actional properties of the verbs from which they are derived. More generally, actionality, or *Aktionsart*, is related to aspect, but refers primarily to "the type of event, specified according to a limited number of relevant properties", while aspect, in its narrow sense, refers to the "specific point of view adopted by the speaker" (Bertinetto 1994, 392). Thus, the latter is more strictly connected to sentence perspective, whereas the former represents semantic properties of verbs as lexical units (cf. Bertinetto 1986, Bertinetto-Delfitto 1992). Therefore, actionality will be extremely relevant for morphological rules. From an actional point of view, one usually distinguishes punctual vs. durative events (e.g., *to fall* vs. *to sleep*), telic or bounded vs. atelic or unbounded events (e.g., *to build* vs. *to smoke*), and static vs. dynamic events (e.g., *to believe* vs. *to run*). The combination of these actional values provides the four fundamental actional classes (states, activities, accomplishments, and achievements, cf. Vendler 1967). It is well known, however, that verbs display different actional values according to the different situational contexts (cf. Smith 1997) in which they occur (namely the presence and the type of verbal arguments, adverbials, etc.). Therefore, it can

sometimes be difficult to establish the actional value of a predicate. Following Brinton (1993), predicates will be classified according to the denoted “prototypical” situation, which can be considered basic. For example, *smoking* is usually an atelic or unbounded activity, although *smoking a cigarette* is a telic or bounded predicate.

In this paper, I will discuss the possible interactions between the morphological rules that form action nouns and the actional content of verbs. In this perspective, an “external” and an “internal” type of actionality must be distinguished. By “external” actionality, the actional value proper of the deverbal noun is meant, which is responsible, for example, for the difference in grammaticality between the following Italian sentences taken from Castelli (1988, 346):

- (1) (a) \*Il dondolio della culla è stato improvviso.  
 ‘The rocking of the cradle was sudden’  
 (b) Il dondolio della culla è durato a lungo.  
 ‘The rocking of the cradle lasted for a long time’

In (1a), the actional value of the deverbal noun is incompatible with the predicate of the sentence *essere improvviso* ‘to be sudden’, in the same way that the adverb *improvvisamente* ‘suddenly’ is incompatible with the process *dondolare* ‘to rock’ (cf. (2a)), while this is not true for (2b):

- (2) (a) \*La culla dondolò improvvisamente (per qualche minuto).  
 ‘The cradle rocked suddenly (for some minutes)’  
 (b) La culla dondolò a lungo.  
 ‘The cradle rocked for a long time’

We can, however, attribute a meaning to the sentences in (1a) and (2a), when the ingressive aspect of the verb is selected:

- (3) (a) Il dondolio (= mettersi a dondolare) della culla è stato improvviso.  
 ‘The (beginning of the) rocking of the cradle was sudden’  
 (b) La culla dondolò (= cominciò a dondolare) improvvisamente.  
 ‘The cradle began rocking suddenly’

External actionality can crucially be dependent on the morphological process forming the action noun, as in the following Dutch examples (cf. Bartsch 1986, 19):

- (4) (a) De verzakking van het huis voltrok zich in twee jaar.  
 'The sinking-in of the house happened in two years'
- (b) ?Het verzakken van het huis voltrok zich in twee jaar.

The unacceptability of (4b) reveals that in Dutch the nominalized infinitive of a verb such as *verzakken* 'to sink-in' is incompatible with a bounded predicate such as *zich voltrekken*, whereas *verzakking* is acceptable in the same context. In this case, two different processes of derivation have given rise to deverbal nouns with different actional properties.

By internal actionality, the actional value of the affix proper is meant, which is revealed by the internal structuring of the deverbal noun, when for instance the actional value of an affix is incompatible with the actional value of a verb. In Thai, for example, two different prefixes select two different actional classes. According to Comrie-Thompson (1985, 351), the prefix *kaan* derives processual deverbal nouns, whereas the prefix *khvam* derives non-processual nouns:

- (5) chyâ → kaan chyâ 'the believing (process.)'  
 chyâ → khvam chyâ 'the belief (non-process.)'

Thai does not possess adjectives; instead, verbs are employed to carry the attributive function. Notice that *kaan* is incompatible with verbs having a stative attributive value, as in (6):

- (6) dii 'good' → khvam/\*kaan dii 'goodness'  
 suäj 'beautiful' → khvam/\*kaan suäj 'beauty'

In this case, the actional value of the prefix *kaan* is incompatible with the actionality of the basic verbs. In other words, *khvam* selects a particular actional class of verbs. Similarly, the Italian suffix *-za*<sup>1</sup> (cf. *distanza*, *convivenza*, *permanenza*, etc.) mostly selects stative verbs as possible inputs, as shown in Gaeta (1999). Moreover, the deverbal nouns formed with the suffix *-za* dis-

<sup>1</sup> There is no space here to discuss the format of the Italian suffix *-za* (for more details cf. Gaeta 1998; 1999). I will assume that the suffix operates on present participles (cf. *distante* → *distanza*, *convivente* → *convivenza*, etc.), with the supplementary addition of a vowel deletion rule which operates everywhere in Italian word formation (cf. Scalise 1983): *distante+za* → *distant∅+za* → *distan[t:s]a* → *distanza* ([nt:s] is morphotactically not allowed). This solution is easier than assuming a suffix *-V-nza*, where V represents the theme vowel of a verb, since this hypothesis cannot cope with cases like *preferire* → *preferenza*, \**preferinza*. Moreover, the format *-za* is assumed by Rainer (1989, 229) for a (semantically very close) suffix producing deadjectival nouns (cf. *elegante* → *eleganza*, *intelligente* → *intelligenza*, etc.).

play a stative external actionality, as they cannot be combined with bounded predicates:

- (7) (a) \*La convivenza di Giovanni con Roberta si compì in due anni.  
 'Giovanni's living-together with Roberta has been completed in two years'.  
 (b) \*La permanenza di Antonio a Roma è stata completata.  
 'Antonio's permanence in Rome has been achieved'.

Therefore, internal actionality expresses the selectional properties of an affix with respect to the base verb, whereas external actionality refers to the actional values proper to an affix that emerge in the derivational process. More generally, we can imagine four types of interactions between internal and external actionality, as sketched in (8):

- (8) (a)  $A_i = \emptyset$   $A_e = \emptyset$  e.g. Du. *het verzakken*  
 (b)  $A_i = \emptyset$ ;  $A_e = F_x$  e.g. Du. *de verzakking*  
 (c)  $A_i = F_x$ ;  $A_e = \emptyset$  e.g. Thai *khwam suǎj*, It. *permanenza*  
 (d)  $A_i = F_x$ ;  $A_e = F_y$  e.g. It. *la nuotata*  
 where  $A_i$  = internal actionality;  $A_e$  = external actionality;  
 F = actional feature.

In the first case (cf. (8a)), the affix does not display selection restrictions or modify the actional value of the verb. For example, the process of nominalization of the Dutch infinitive does not have any relevant effect on the actionality of an atelic verb, as seen in (4) above. On the other hand, the Dutch suffix *-ing* causes the telicization of an unbounded basic predicate. As represented in (8b), the external actionality of this suffix must be specified with the feature [+ bounded].

In the third case (cf. (8c)), the internal actionality of the affix is able to select a particular actional class of verbs, without modifying the external actionality of the nominalized predicate; accordingly, the deverbal noun preserves the selected internal actionality, as seen in (6) and in (7) above.

Finally (cf. (8d)), we can imagine that an affix displays selection restrictions and at the same time forms deverbal nouns that have a particular actional value, which is different from that of the basic verb. In what follows, I will try to show that this is the case for an Italian suffix which forms a consistent number of deverbal and denominal nouns and is quite productive (cf. Scalise 1983, Gatti-Togni 1991, Mayo et al. 1995):

- (9) (a) dormire 'to sleep' → dormita  
 leggere 'to read' → letta  
 mangiare 'to eat' → mangiata  
 scorrere 'to run through' → scorsa
- (b) gomito 'elbow' → gomitata  
 asino 'donkey' → asinata  
 giorno 'day' → giornata  
 sedia 'chair' → sedziata

Diachronically, the suffix *-ata* that is used to form denominal nouns (cf. *gomitata*, *asin-ata*, etc.) corresponds to the feminine form of the past participle, as can be seen in the case of irregular verbs such as *leggere* 'to read' → *letta*, *scorrere* 'to run through' → *scorsa*, where the deverbal nouns are directly derived on the basis of the respective past participle (cf. *letto*, *scorso*). However, because of the huge number of regular verbs of the I inflectional class ending in *-are* such as *mangiare* → *mangiata*, the ending *-ata* was successively extended to nouns, giving rise to forms like those reported in (9b) (cf. Tekavčić 1972, 57). In what follows, I will not go into further formal problems. My purpose is to investigate the semantic content of the word formation rule forming deverbal nouns, yet I will not provide a precise format for the word formation rule. In the course of the paper, I will speak of ATA-nominals, referring generically to deverbal nouns derived from the feminine form of the past participle.

## 2. ATA-nominals and the "packaging" of information

With respect to other Italian nominalizations, "derivations in *-at(a)* are semantically restricted in such a way that they cannot normally be interpreted as types of actions, but only as individual or instantiated events" (cf. Mayo et al. 1995, 912). Therefore, ATA-nominals cannot be accompanied by the definite article in the generic meaning or by the null article, as is seen in (10):

- (10) (a) Il nuoto/\*la nuotata in piscina rilassa i muscoli.  
 'Swimming in the pool relaxes muscles'
- (b) Domani ci sarà una gara di nuoto/\*nuotata.  
 'A swimming competition will take place tomorrow'

*Nuotata* cannot denote the event or the process as such, but only a single instantiation of it. In fact, (10a) becomes grammatical, if *nuotata* is modified by a restrictive relative clause, as in (11):

- (11) *La nuotata che faccio di solito rilassa i muscoli.*<sup>2</sup>  
 ‘The swim that I usually have relaxes my muscles’

In Talmy’s (1988, 176) words, “a single instance of the specified equivalent units is taken and set in the foreground of attention”, just as in the English examples of the kind *to breathe* → *take a breath* in the verbal domain, and *furniture* → *a piece of furniture* in the nominal domain. This is the well-known phenomenon of the packaging of the information contained in the predicate (cf. Paprotté 1988; Jackendoff 1991; Brinton 1993). More generally, we can assume a parallel between the verbal and the nominal domains. Bounded predicates can be compared to countable nouns, since they “can be directly or intrinsically counted” (cf. Mourelatos 1978, 429f). On the other hand, unbounded predicates can be compared to mass nouns, since they are subdivisible to infinity. The process of packaging allows one to express a mass noun or an unbounded predicate as a single instantiation respectively of the uncountable entity or of the process. This means that every portion of *sleeping* corresponds to the activity of sleeping, as well as every portion of *water* is still water. On the other hand, a portion of the event of arriving in Budapest cannot be considered the event of arriving in Budapest, as well as a portion of an apple, for instance its core, cannot be directly considered an apple. Thus, the process of packaging, and its opposite, i.e., the process of grinding (cf. Paprotté 1988; Jackendoff 1991; Brinton 1993), allows one to represent predicates and things in the inverse form with respect to their basic properties.

My claim is that ATA-nominals achieve the operation of packaging in the verbal domain. Unbounded predicates are transformed into single and bounded portions of the denoted activity. For this reason ATA-nominals cannot be accompanied by the definite article in the generic meaning; as single instantiations of the relevant activity, they cannot be used to refer to the process as such. Yet if in (10a) the ATA-nominal is accompanied by the indefinite article, the sentence becomes grammatical, as in (12):

- (12) *Una nuotata in piscina rilassa i muscoli.*  
 ‘Swimming in the pool relaxes muscles’

<sup>2</sup> Similar considerations hold true when the ATA-nominal is used in a generic sense as in the following example:

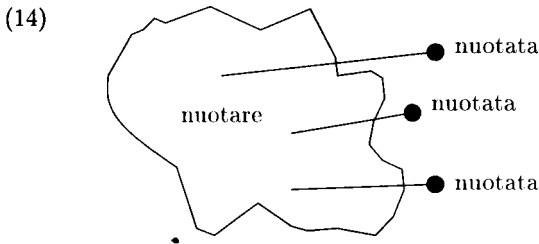
- (i) *La nuotatina giornaliera che mi ha consigliato il medico mi ha fatto bene.*  
 ‘The daily swim the doctor recommended to me was healthy’.

In this respect, Mayo et al. (1995, 912) observe that “this arises from an independent process of generalization that can be applied to any nominal concept”.

As is well known, the indefinite article is usually employed to achieve the operation of packaging in the nominal domain, rendering mass nouns countable as in (13):

- (13) il caffè            'the coffee'    vs.    un caffè            'a coffee'  
 l'acqua            'the water'    vs.    un'acqua        'a water'  
 la birra            'the beer'     vs.    una birra        'a beer'

The mass nouns are thus packaged into single instantiations, as is revealed by a sentence such as *Giovanni ha pagato un caffè, un'acqua e una birra* 'John has paid for a coffee, a water and a beer'. We can represent the operation of packaging fulfilled by the ATA-nominal by the picture in (14), in which single portions of the process denoted by a predicate without boundaries like *nuotare* 'to swim' are extracted and represented as countable points:



To describe this process in a more formal way, we can assume Jackendoff's (1991) ELT (i.e., 'element of') function. This function maps its argument onto a subentity of the larger entity denoted by the argument, as seen in (15):

$$(15) \text{ nuotata} = \left[ \text{ELT} \left( \begin{array}{c} +b, -i \\ \left[ \begin{array}{c} -b, +i \\ \text{SWIM} \end{array} \right] \end{array} \right) \right]$$

From the process of swimming, which is unbounded (i.e.,  $-b$ ), but with an internal structure (i.e.,  $+i$ ), the ATA-nominal extracts a subentity provided with the opposite features, (i.e.,  $+b, -i$ ). Given the function of packaging operator achieved by ATA-nominals, it is not surprising that the selected predicate is durative and dynamic, i.e., a process, from which a single portion can be extracted. Thus, stative verbs cannot be the input of a packaging operator:<sup>3</sup>

<sup>3</sup> Elsewhere (cf. Gaeta 1999), it has been shown that in Italian predicates provided with the actional feature  $[-\text{dynamic}]$  are compatible in a productive way only with the suffix *-za* (cf. *distanza*, *permanenza*, etc.).

- |      |           |              |   |             |
|------|-----------|--------------|---|-------------|
| (16) | conoscere | 'to know'    | → | *conosciuta |
|      | credere   | 'to believe' | → | *creduta    |
|      | giacere   | 'to lie'     | → | *giaciuta   |

Before going into other selection restrictions of ATA-nominals, it is necessary to investigate another important property of them: they form a periphrasis with the support verbs *dare* and *fare*.

### 3. The periphrases *fare/dare una V-ATA*

Similar to English constructions like *to breathe* → *take a breath*, Italian displays periphrases involving the support verbs<sup>4</sup> *dare* 'to give' and *fare* 'to do', which bear the grammatical features of tense, mood, person, etc., whereas the ATA-nominals contain the lexical meaning of the verb of the basic sentence (cf. Salvi 1988, 79ff):

- (17) (a) I bambini fanno una dormita.  
'The children are having a sleep'
- (b) La mamma dà un'ordinata alla casa.  
'The mother puts the house in order'

The periphrasis with *fare* 'to do' mostly takes ATA-nominals derived from intransitive verbs<sup>5</sup> (cf. (17a)), whereas the periphrasis with *dare* 'to give' usually takes ATA-nominals derived from transitive verbs (cf. (17b)). According to Salvi (1988, 81), with respect to basic sentences both periphrases generally represent the event as short and occasional. Let us now investigate the Italian periphrases more deeply, by taking advantage of the very precise analysis on the similar English constructions with the support verbs *have* (or *take*)<sup>6</sup>

<sup>4</sup> For the notion of support verb, first elaborated for French, see, among others, Gross (1981), and, for Old Italian, La Fauci (1979).

<sup>5</sup> However, transitive verbs may also form periphrases with *fare*, e.g., *fare una mangiata* from *mangiare* 'to eat', *fare una bevuta* from *bere* 'to drink', etc. As will be shown below (cf. (25)), the relevant property here seems to be the actional characteristic of the predicates, since only activities are compatible with the periphrasis (cf. *fare una mangiata di pizza* vs. \**fare una mangiata della pizza*).

<sup>6</sup> The difference between the two support verbs, which correspond to Italian *fare*, is partly structurally motivated (cf. Wierzbicka 1988, 337ff), partly dialectal: British (and Australian) English prefers *have*, whereas American English makes use of *take* (cf. Dixon 1991, 338).



and *give* accompanied by a converted deverbal noun as in (18), conducted by Wierzbicka (1988) and Dixon (1991, 336ff):

- (18) John had a walk / a swim / a lie-down.  
 She gave me a push / a kiss / a look.

Notice that the English periphrases share more or less the same distribution as the Italian one, since *have* is preferably (but not only, cf. *have a lick of the ice cream*) combined with intransitive verbs, whereas *give* usually takes transitive verbs. Moreover, the global semantics of the English periphrasis is similar to the Italian one, since—according to Wierzbicka (1988, 297)—it “presents the action (or the process) as limited in time”. In other words, the periphrastic construction portrays the event as short and occasional. To distinguish between true periphrasis and other similar constructions, Dixon (1991, 339ff) adopts the four criteria listed in (19):

- (19) (a) form: a periphrastic construction must show  
 (i) the same subject as the basic sentence;  
 (ii) *have, take* or *give* as the main verb;  
 (iii) the base form of the verb of the basic sentence as head of a post-predicate NP, preceded by the indefinite article *a ~ an*;
- (b) meaning: the periphrastic sentence should have essentially the same meaning as the basic sentence;
- (c) adverb/adjective correspondence: the way in which an adjective provides semantic modification to the head of an NP is similar to the way in which an adverb modifies a verb, like in *Mary kissed him passionately* → *Mary gave him a passionate kiss*.
- (d) preservation of peripheral constituents: all peripheral constituents of the basic sentence should be exactly preserved in the periphrastic construction like in *I always swim in the pool before breakfast on weekdays* → *I always have a swim in the pool before breakfast on weekdays*.

Space prevents me from describing the English periphrasis more thoroughly. In what follows, I will try to apply Dixon’s criteria to the Italian periphrastic construction in order to distinguish it from other similar constructions, and, above all, to establish its semantic value. However, the main focus of the investigation will remain on ATA-nominals, which occupy the place of the converted deverbal nouns in the English periphrasis; we will see that compatibility with the periphrastic construction is an important test to distinguish between productive and lexicalised derivatives.

In this perspective, notice that the periphrastic construction allows us to distinguish it from the case in which an ATA-nominal is fully lexicalised, as in (20):

- (20) (a) *Alla festa abbiamo fatto una ballata e siamo subito andati via.*  
 'At the party, we had a dance and went suddenly away'
- (b) *In ricordo della donna amata, il poeta fece una ballata molto commovente.*  
 'In memory of his lover, the poet composed a very touching ballad'

In (20b), *ballata* has a concrete value, i.e., 'poem'; hence the sentence does not contain the periphrastic construction (apart from the obvious pun!). In fact, in (20b) the criteria seen in (19) above are violated, since (20b) does not presuppose a basic sentence such as *In ricordo della donna amata, il poeta ballò in maniera molto commovente* 'In memory of his lover, the poet danced in a very touching way'.<sup>7</sup> It can easily be checked how the above criteria hold true for the sentence in (20a), where we find the periphrastic construction.

Apart from this rather easy case, however, where the ATA-nominal is clearly lexicalised, the criteria in (19) help us discriminate in much more complex sentences. For example, criterion (19c), i.e., the adverb/adjective correspondence, distinguishes between the case in which we find the periphrasis and the case in which the nominal is lexicalised:

- (21) (a) *Gli alpinisti fecero una discesa rapida verso il paese.*  
 'The mountaineers made a quick descent to the country'
- (b) *Gli alpinisiti fecero una discesa ripida verso il paese.*  
 'The mountaineers made a steep descent to the country'

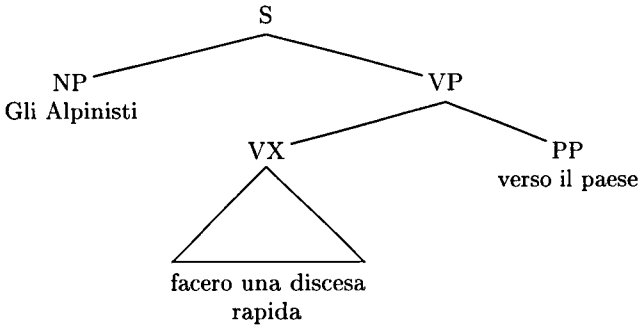
In (21b) the adjective *ripida* 'steep' does not correspond to the adverb of the basic sentence (cf. *\*Gli alpinisiti discesero ripidamente verso il paese* 'The mountaineers descended steeply to the country'), but refers to the path followed by the mountaineers, whereas *rapida* 'quick' in (21a) modifies as an adverb the predicate in the basic sentence *Gli alpinisiti discesero rapidamente al paese* 'The mountaineers quickly descended to the country'. Finally, (21b) also violates criterion (19d) above, i.e., the preservation of peripheral constituents. In fact, the modifier *verso il paese* 'to the country' qualifies the

<sup>7</sup> Notice that in the case of the lexicalised ATA-nominal (cf. (20b) above), the inflected verb can be substituted by a synonym, which is sometimes stylistically preferred as in (ii) below, while this is not the case with the true periphrastic construction:

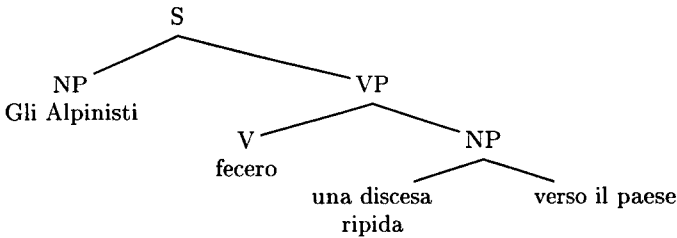
- (i) *\*Alla festa abbiamo compiuto una ballata e siamo subito andati via.*  
 (ii) *Il poeta compose una ballata molto commovente.*

concrete head noun *discesa*, by attributing a property (i.e., that of being directed to the country) to it, rather than being the goal argument of the basic predicate *discendere*. In terms of (really simplified) syntactic structure, this difference can be represented as in (22), where (22a) corresponds to (21a), and (22b) corresponds to (21b):

(22) (a)



(b)



where VX = complex verb

Another signal of a certain degree of lexicalization of the ATA-nominal is the combinability with the definite article. In fact, in (21b) it is possible to employ the definite article to modify *discesa*, whereas in (21a) this is excluded:

- (23) (a) Gli alpinisti fecero la discesa rapida verso il paese.  
 'The mountaineers made the quick descent to the country'
- (b) Gli alpinisti fecero la discesa ripida verso il paese.  
 'The mountaineers made the steep descent to the country'

The definite article in (23a) forces the same interpretation as in (23b), in which *discesa* is lexicalised. Therefore, we can say that *discesa* presents the ambiguity of a word undergoing a process of lexicalization. The employment of the definite article instead of the indefinite one is a good criterion to test the semantics of the periphrastic construction. Consider the sentences in (24):

- (24) (a) Carlo ha raccolto materiale da riciclare.  
 'Charles has gathered recyclable material'  
 (b) Carlo ha fatto la raccolta di materiale da riciclare.  
 (c) Carlo ha fatto una raccolta di materiale da riciclare.

With respect to the basic sentence in (24a), the sentence in (24c), containing the periphrastic construction, represents the event as rather approximate and imprecise; its external boundary is not well-defined. In fact, from (24c)—as opposed to (24b)—we do not obtain the information that Charles has gathered all of the recyclable material, but that he has only done a very imprecise job. Notice that in this way the object of the basic verb is backgrounded; what is represented as relevant by the periphrastic construction is the approximate way in which the subject participates in the process. The backgrounding of the object is the reason why it is impossible to have the periphrastic construction with a predicate denoting an activity with a precise *télos*:

- (25) (a) Maria ha fatto una mangiata di pizza.  
 'Mary stuffed herself with pizza'  
 (b) \*Maria ha fatto una mangiata della pizza.  
 'Mary stuffed herself with the pizza'

This also holds true for the periphrastic construction containing *dare*. The activity is represented as imprecise and unbounded:

- (26) \*Sara diede una piegata ai vestiti in due ore.  
 'Sarah folded up dresses in two hours'  
 \*Antonio ha dato una pettinata a Mario in due ore.  
 'Tony combed Mario's hair in two hours'

Thus, the periphrastic construction as a whole achieves a detelicizing function: it selects an activity and represents it as short and occasional, in which—as Dixon (1991, 346) notes—"the subject indulges ... for a certain period". In other words, the periphrastic construction represents the activity as subject-oriented. In fact, the subject of the periphrasis must be human or intentional as shown by the following sentences:

- (27) (a) ??La Gioconda ha fatto un'attesa di dieci anni prima di essere restaurata.  
 'The Gioconda has awaited ten years before being restored'  
 Tina ha fatto un'attesa di due ore prima di essere ricevuta.  
 'Tina has awaited two hours before being admitted'

- (b) ??La pioggia diede una riempita alla piscina in giardino.  
 'The rain filled up the swimming pool in the garden'

Il giardiniere diede una riempita alla piscina in giardino.  
 'The gardener filled up the swimming pool in the garden'

Bounded predicates, i.e., in Vendlerian terms accomplishments and achievements, are excluded from the periphrastic construction and, accordingly, do not form an ATA-nominal, as shown by the ungrammaticality of (28):

- (28) \*fare una costruita/\*fare un'arrivata/\*fare una partita, etc.  
*lit.* 'to make a build, an arrive, a leave'  
 \*dare una costruita alla casa/\*dare un'uccisa al gangster, etc.  
*lit.* 'to give a build to the house, a kill to the gangster'

In Wierzbicka's (1988, 323) words, these predicates are excluded from the periphrastic construction since they have "an external goal and a natural boundary (reached when the goal is attained)". The case of *ammazzata* from *ammazzare* 'to slaughter' is very interesting, since it actually occurs as an ATA-nominal in the periphrastic construction:

- (29) \*Il macellaio fece un'ammazzata di maiali per preparare le salsicce.  
 'The slaughterer slaughtered up pigs to make sausages'  
 Il macellaio fece un'ammazzata per preparare le salsicce.  
 'The slaughterer worked hard to make sausages'

However, *ammazzata* does not refer to the event of slaughtering, but to the slaughterer's getting tired during his work.

#### 4. Marginal cases, exceptions, and the productivity of -ATA

Having illustrated the basic semantics of ATA-nominals, let us now try to extend the boundaries of our investigation. There is indeed a number of verbs that do not fit into the picture sketched above. They form (roughly) two classes. First, there is a number of achievements that display an ATA-nominal which is compatible with the periphrastic construction:

- |      |                          |                              |                            |
|------|--------------------------|------------------------------|----------------------------|
| (30) | caduta (< 'to fall')     | entrata (< 'to go in')       | salita (< 'to go up')      |
|      | calata (< 'to go down')  | fermata (< 'to stop')        | scappata (< 'to rush off') |
|      | cascata (< 'to fall')    | rientrata (< 'to go backin') | scivolata (< 'to slip')    |
|      | comparsa (< 'to appear') | risalita (< 'to go back to') | uscita (< 'to go out')     |
|      | discesa (< 'to go down') | riuscita (< 'to go out')     | venuta (< 'to come')       |

There is no space to analyse the exact behaviour of these verbs (for a thorough analysis, cf. Gaeta 1998). Let me just observe that many of these derivatives are partially or fully lexicalised, since they violate the criteria established in (19) above. For example, *entrata* and *venuta* are possible only in constructions like the ones in (31a); otherwise (cf. (31b)), they are ungrammatical:

- (31) (a) fare una entrata improvvisa  
 'to make a sudden entrance'  
 fare una venuta improvvisa  
 'to make a sudden coming'
- (b) \*fare un'entrata nella stanza  
 'to make an entrance into the room'  
 ??fare una venuta da Roma  
 'to make a coming from Rome'

Thus, ATA-nominals formed on the basis of achievements that are compatible with the periphrastic construction seem to be rather marginal with respect to the large set of words derived from unbounded verbs.<sup>8</sup>

Second, there is a group of verbs that are incompatible with the periphrastic construction:

<sup>8</sup> Many verbs listed in (30) belong to the so-called class of ergative (or unaccusative) verbs (cf. Salvi 1988, 47ff). Drawing on this aspect, Bordelois (1993, 172) has claimed that in Spanish "el sufijo *-Da* ... se extiende a todos los ergativos, en sentido estricto o en sentido laxo". The Spanish suffix *-Da*, which corresponds to It. *-ATA*, since it also derives deverbal nouns from the feminine past participle form, would only be compatible with ergative verbs, either of the strict type (cf. *caída* 'fall', *salida* 'exit', *llegada* 'arrival', etc.), or of the loose type (i.e., in her view, verbs having a non-agentive subject, such as *vista* 'sight', *mirada* 'look', etc.). However, this analysis cannot cope with a number of activities, which, as well as in Italian, give rise to deverbal nouns such as *chupada* from *chupar* 'to suck up', *mamada* from *mamar* 'to suck', *pisada* from *pisar* 'to trample on', etc. (cf. Rainer 1993, 438ff; Lüdtke 1978, 363ff). Moreover, it seems that, at least in South American Spanish, "in Verbindung mit Funktionsverben wie *dar*, *echar*, *pegar*, u.a. kann ... in der Umgangssprache fast jedes Verb (der ersten Konjugation) durch *-da* in ein Nomen Actionis verwandelt werden: *dar cabeceadas*, *pegar una calentada*,  *darse una desorientada*, *dar una hablada a alguien*, *echarse una buena investigada*, *echar una platicada*, usw." (Rainer 1993, 440).

|      |                              |                              |
|------|------------------------------|------------------------------|
| (32) | aggiunta (< 'to add')        | promessa (< 'to promise')    |
|      | andata (< 'to go')           | proposta (< 'to propose')    |
|      | cacciata (< 'to chase away') | richiesta (< 'to ask for')   |
|      | difesa (< 'to defend')       | rimessa (< 'to put back')    |
|      | durata (< 'to last')         | ripresa (< 'to recover')     |
|      | messa (< 'to put')           | rotta (< 'to break')         |
|      | offerta (< 'to offer')       | scomparsa (< 'to disappear') |
|      | offesa (< 'to offend')       | scoperta (< 'to discover')   |
|      | presa (< 'to take')          | spesa (< 'to spend')         |
|      | pretesa (< 'to pretend')     |                              |

For the nominals in (32), it is easy to verify the incompatibility with the periphrastic construction, by applying the criteria in (19) above. What is relevant for our purposes is that all these nominals, listed nearly exhaustively in (30) and (32), do not usually behave as packaged pieces of the information contained in the predicate, although they may be compatible with the support verbs *fare* and *dare*.<sup>9</sup> In fact, they are true action nouns, as is shown in their blocking other possible deverbal nouns:

|          |          |                             |
|----------|----------|-----------------------------|
| (33) (a) | caduta   | *cadimento/*cadizione       |
|          | entrata  | *entramento/*entratazione   |
|          | fermata  | *fermamento/*fermatazione   |
|          | venuta   | *venimento/*venizione       |
| (b)      | cacciata | *cacciamento/*cacciatazione |
|          | difesa   | *difendimento/*difensione   |
|          | offesa   | *offendimento/*offensione   |
|          | promessa | *promettimento/*promissione |

Moreover, they are usually rather old: none of them is posterior to the sixteenth century. In other words, they constitute a sub-class within the ATA-nominals that cannot be considered central for determining the role of the suffix within the system. Indirectly, this confirms the importance of the periphrastic construction as a test for the investigation of ATA-nominals; only derivatives compatible with the periphrasis display the function of packaging operator.

<sup>9</sup> The obtained periphrases violate the criteria discussed in (19) above. For example, the following sentences do not correspond either semantically or syntactically to the respective matrix sentences \**Il professore aggiunse lungamente critiche* and \**Il miliardario offrì cospicuamente denaro*:

- (i) Il professore fece una lunga aggiunta di critiche.  
'The professor made a long addition of criticisms'
- (ii) Il miliardario fece una cospicua offerta di denaro.  
'The millionaire made a conspicuous offer of money'

In striking contrast to these lexicalised cases, ATA-nominals firstly attested in this century display very different properties. Recall that the suffix is quite productive. On the basis of *DISC*, I counted about 100 new deverbal derivatives only in this century, from which the ones in (34) are taken:

|      |                                                |      |
|------|------------------------------------------------|------|
| (34) | accelerata (< accelerare 'to accelerate')      | 1967 |
|      | aggiustata (< aggiustare 'to adjust')          | 1954 |
|      | foraggiata (< foraggiare 'to fodder')          | 1970 |
|      | insaponata (< insaponare 'to soap')            | 1936 |
|      | litigata (< litigare 'to quarrel')             | 1932 |
|      | ordinata (< ordinare 'to put in order')        | 1939 |
|      | ospitata (< ospitare 'to give hospitality to') | 1997 |
|      | regolata (< regolare 'to regulate')            | 1978 |
|      | rimodernata (< rimodernare 'to modernize')     | 1991 |
|      | riscaldata (< riscaldare 'to warm')            | 1940 |
|      | scrutata (< scrutare 'to scrutinize')          | 1960 |
|      | stirata (< stirare 'to iron')                  | 1960 |

Moreover, it is possible to form ATA-nominals from borrowed verbs like the following:

|      |                            |   |           |
|------|----------------------------|---|-----------|
| (35) | dribblare (< 'to dribble') | → | dribblata |
|      | slurpare (< 'to slurp')    | → | slurpata  |
|      | sniffare (< 'to sniff')    | → | sniffata  |
|      | zumare (< 'to zoom')       | → | zumata    |

They all match the type I have tried to sketch up until now. Namely, they are derived from unbounded dynamic verbs, function as a packaging operator, and are compatible with the periphrastic construction. Moreover, they usually do not give rise to lexical blocking with respect to the true action nouns:

|      |             |                |
|------|-------------|----------------|
| (36) | accelerata  | accelerazione  |
|      | aggiustata  | aggiustamento  |
|      | ammazzata   | ammazzamento   |
|      | foraggiata  | foraggiamento  |
|      | insaponata  | insaponamento  |
|      | litigata    | litigio        |
|      | nuotata     | nuoto          |
|      | regolata    | regolamento    |
|      | rimodernata | rimodernamento |
|      | riscaldata  | riscaldamento  |
|      | sbrodolata  | sbrodolamento  |
|      | scrutata    | scrutamento    |
|      | stirata     | stiramento     |
|      | verniciata  | verniciatura   |
|      | vuotata     | vuotamento     |



Therefore, it seems that ATA-nominals contain words of rather different properties; there is a central core formed by derivatives that achieve the function of packaging operator. Predicates selected by the suffix -ATA must be dynamic and durative. This productive core is surrounded by derivatives that are lexicalised since they assume the value of true action nouns. This sub-class is not well-defined, nor can it be freely enlarged. Moreover, it is extremely small compared to the productive core.

### 5. The denominal ATA-nominals

At the beginning of this paper, I mentioned that besides ATA-nominals there is a large number of derivatives from nominal bases. According to their specific meanings, they can be grouped as follows (cf. Scalise 1983; Gatti-Togni 1991; Mayo et al. 1995):

- |                             |                                                                   |
|-----------------------------|-------------------------------------------------------------------|
| (37) (a) blow of N          | gomito 'elbow' → gomitata<br>sedia 'chair' → sediatà              |
| (b) action typical of N     | asino 'donkey' → asinata<br>pagliaccio 'clown' → pagliacciata     |
| (c) quantity contained in N | cucchiàio 'spoon' → cucchiàiatà<br>secchiò 'bucket' → secchiàtatà |
| (d) augmentative of N       | cancellò 'gate' → cancellatà<br>vallè 'valley' → vallatà          |
| (e) period of time N        | giornò 'day' → giornatà<br>mattinà 'morning' → mattinatà          |

From a diachronic point of view, the suffix forming denominal ATA-nominals has clearly originated from a reanalysis of the deverbal nouns. As Rohlfs (1969, 444) puts it, "per il fatto che per esempio *ventata* poteva appartenere tanto a *ventare*, quanto esser derivata direttamente da *vento*, divenne possibile ottenere anche da sostantivi dei derivati simili: *occhiata*, *bambinata*, *birbonata*". The selection of the -ata-variant is thus due to the high number of verbs belonging to the -are inflectional class, from which ATA-nominals are derived. Notice that among the several meanings listed in (37) above, the types 'augmentative of N' (cf. (37d)) and 'period of time N' (cf. (37e)) are restricted to a limited set of words, which cannot be freely extended, as is shown by the following examples:

- (38) (a) collina 'hill' → \*collinata  
 pianura 'plain' → \*pianurata
- (b) pomeriggio 'afternoon' → \*pomeriggiata  
 settimana 'week' → \*settimanata

Thus, they seem to be no longer productive, and I will exclude them from the rest of the discussion. As has been observed by Simone (1993, 51), the nominals in (37a–c) share the basic framework meaning, according to which the nominal denotes a single instantiation of the action in which the considered substance is involved. Schematically, one can reduce the global semantics of these nominals to the basic meaning 'a single, short and fast instantiation of an action carried by/typical of X'. Thus, a *gomitata* is a single instantiation of an action carried by an elbow (i.e., that of hitting someone), as well as an *asinata* is a single instantiation of an action typical of a donkey (i.e., metaphorically, that of acting like a lazy schoolboy). Presumably, the difference between (37a) and (37b) can be accounted for by the feature [+ human], which defines the set of bases selected by (37b) (cf. *bambino* 'child' → *bambinata*, *stupido* 'silly' → *stupidata*, etc.), whereas this is not the case for (37a) (cf. *ginocchio* 'knee' → *ginocchiata*, *martello* 'hammer' → *martellata*, etc.). (37c) also fits into this schema, if one considers that *secchiata* does not simply mean the quantity of water contained in a bucket. Much more, it denotes a single instantiation of the action in which a typical container such as a bucket is involved. Thus, *secchiata* can refer either to the action of blowing or of being blown by a bucket, or to the quantity of water contained in a bucket that can be poured onto someone:

- (39) (a) Gianni ha preso una secchiata sulla testa.  
 'G. was hit on the head with a bucket'
- (b) Gianni ha preso una secchiata d'acqua sulla testa.  
 'G. was hit on the head with a bucket of water'

Since a bucket is a typical container, *secchiata* can then be extended to denote the quantity of substance a bucket can contain. Notice, however, that the rough and imprecise character of the basic meaning is still present in the shifted meaning, as is shown by the following sentences (cf. Samek-Ludovici 1997):

- (40) (a) Questa vasca contiene esattamente 23 litri d'acqua.  
 'This basin contains exactly 23 liters of water'
- Un ettolitro d'acqua sono esattamente 100 litri d'acqua.  
 'One hectolitre of water is exactly 100 liters of water'

(b) ??Questa vasca contiene esattamente 7 secchiate d'acqua.  
'This basin contains exactly 7 buckets of water'

??Una secchiata d'acqua sono esattamente 2 litri d'acqua.  
'A bucket of water is exactly 2 liters of water'

Yet, the semantic shift 'blow of N' → 'quantity contained in N' is not automatic. There are nominals formed on the basis of typical containers such as *bottigliata* from *bottiglia* 'bottle', or *padellata* from *padella* 'pan', in which the shifted meaning is not attested. In these cases, the nominal only displays the basic meaning 'blow of N'.

Thus, denominal ATA-nominals share with the deverbal ATA-nominals the basic meaning 'single instantiation of an action carried out in a rather short and imprecise way'. Notice that denominal ATA-nominals are also compatible with the periphrastic constructions containing a support verb (cf. *fare una bambinata*, *dare una secchiata*, etc.). In certain cases (e.g., *martellata* 'blow of hammer'), it would be arbitrary to decide whether the ATA-nominal is derived from the substantive (i.e., *martello* 'hammer') or from the verb (i.e., *martellare* 'to hammer'). However, in the standard analysis adopted by Scalise (1983), the striking semantic similarities of the denominal and of the deverbal ATA-nominals are not accounted for. In order to rescue Aronoff's (1976) Unitary Base Hypothesis, Scalise (1983) proposes two different derivational rules containing two different suffixes, in which the first one deals with verbs and the second one with nouns:

- (41) (a)  $[V]_{\text{PASTPART}} + [-a]_{\text{SUFF}} \rightarrow [V]_{\text{NFEM}}$   
 $[\text{mangiato}]_V + [-a]_{\text{SUFF}} \rightarrow [\text{mangiat-}\emptyset]_V + [-a]_{\text{SUFF}} \rightarrow [\text{mangiata}]_{\text{NFEM}}$
- (b)  $[N] + [-ata]_{\text{SUFF}} \rightarrow [N]_{\text{NFEM}}$   
 $[\text{gomito}]_N + [-ata]_{\text{SUFF}} \rightarrow [\text{gomit-}\emptyset]_N + [-ata]_{\text{SUFF}} \rightarrow [\text{gomitata}]_{\text{NFEM}}$

Independent of the inconsistency of assuming an inflectional suffix to form deverbal nouns in (41a) (cf. Thornton 1990), these rules do not express the similar semantic meaning we have observed previously. A unitary treatment has been recently proposed to account for both deverbal and denominal ATA-nominals (cf. Samek-Ludovici 1997).<sup>10</sup> In particular, this author has suggested treating the denominal nominals as derived by means of the rule in (41a) applied on a non-surfacing intermediate predicate, a so-called "ghost verb", derived by conversion from the basic noun:

<sup>10</sup> For a similar unitary treatment (albeit couched in a Coserian approach) of the corresponding Spanish and French derivatives, see Dietrich (1994).

- (42) gomito → \*gomitare → gomitata  
 asino → \*asinare → asinata

In Samek-Ludovici's (1997) account, postulating an intermediate false step is needed to explain the verbal properties possessed by denominal ATA-nominals. In the periphrastic construction with the support verb *dare*, the argument structure of the intermediate predicate \**gomitare* would filter into the support verb as in the case of the existing verb *lavare* 'to wash' with respect to *lavata* 'wash':

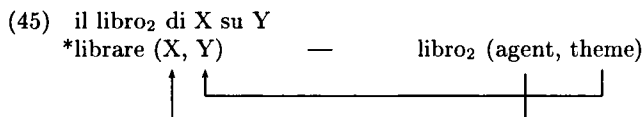
- (43) (a) X dà una lavata a Y  
 dare (X, Y) — lavata (agent, theme)
- 
- (b) X dà una gomitata a Y  
 \*gomitare (X, Y) — gomitata (agent, theme)
- 
- The diagrams show two examples of argument structure inheritance. In (a), the verb 'dare' has arguments X and Y, and the noun 'lavata' has an agent and a theme. A horizontal line connects 'lavata' to 'dare', with a vertical line from 'lavata' to the theme position and another from 'dare' to the agent position. In (b), the verb '\*gomitare' has arguments X and Y, and the noun 'gomitata' has an agent and a theme. A horizontal line connects 'gomitata' to '\*gomitare', with a vertical line from 'gomitata' to the theme position and another from '\*gomitare' to the agent position.

Thus, the ATA-nominal would inherit its argument structure from the existing, but non-surfacing, ghost verb. However, this solution runs into troubles because of its abstractness. Postulating an intermediate false step can be wildly extended to many cases, in which a denominal noun is provided with verbal properties such as an argument structure. In fact, nothing prevents us from assuming an intermediate verbal step for agent nouns derived from nouns, as in the following examples:

- (44) (a) il mercante di fiori                      mercante < \*mercare<sup>11</sup> < merce  
 'the merchant of flowers'
- (b) l'autista di taxi                                autista < \*autare < auto  
 'the taxi-driver'

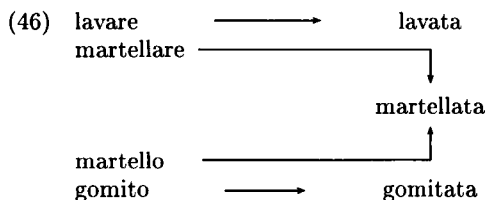
Moreover, the postulation of an underlying ghost verb can be freely extended to any case of nouns displaying argument structure, as in *il libro di Sartre su Flaubert* 'Sartre's book on Flaubert', in which a ghost verb \**librare*, derived from a basic *libro*<sub>1</sub>, can theoretically be postulated to account for the argument structure of *libro*<sub>2</sub>:

<sup>11</sup> Notice that the verb *mercare* is actually attested for Old Italian, but has now completely disappeared (cf. *DISC*, s.v.).



Finally, the ghost verb hypothesis runs against the intuition of speakers, who, at least in some varieties of Italian, do make use of verbs corresponding to a denominal ATA-nominal, such as *gomitare*<sup>12</sup> ‘to push with elbows’, reported by some Italian dictionaries, or *occhiare* ‘to look’ from *occhio* ‘eye’, heard in Rome. These cases, however, are back-formations with respect to the bases *gomitata* and *occhiata*, rather than the surfacing of the postulated ghost verb, as can be shown by applying Marchand’s (1964) criterion of semantic implication.

In the absence of a convincing motivation for postulating a huge number of ghost verbs, this hypothesis must be discarded. Thus, if we want to account for the striking similarities observed earlier, we are left with a solution that violates the Unitary Base Hypothesis.<sup>13</sup> However, the picture of ATA-nominals appears to be semantically very coherent, so that in several cases we can speak of a double motivation (cf. Szymanek 1988, 65):



This representation also takes into consideration the diachronic origin of the suffix *-ata*, which developed, as mentioned above, from triplets such as *vento*

<sup>12</sup> The actual verb in Standard Italian denoting the action of pushing with elbows is the parasynthetic *sgomitare*, formed by means of the so-called intensifying prefix *s-* (for the whole question of parasynthesis in Italian, cf. the different positions expressed by Scalise (1983) and Crocco Galéas-Iacobini (1993)).

<sup>13</sup> Notice, however, that the case of word formation rules selecting different lexical bases is anything but unusual. Among others, the German suffix *-ei* (cf. Fleischer-Barz 1992, 149f) can be mentioned, which forms denominal (cf. *Abtei, Gärtnerrei*) and deverbal nouns (cf. *Aufschneiderrei, Tanzerei*); or the Latin suffix *-ul/-ol-* (cf. Plank 1981, 44), which forms diminutives from nominal (cf. *regulus, filiulus*), from adjectival (cf. *aureolus, blandulus*) and from verbal bases (cf. *misculare*). The Unitary Base Hypothesis has been recently questioned on the basis of semantic arguments by Plag (1998, 237), who puts forward the very interesting idea that “the syntactic category of potential base words is only a by-product of the semantics of the process”.

'wind' – *ventare* – *ventata*, where a kind of rule telescoping seems to have taken place after (or because of) the disappearance of the intermediate verb.

## 6. Conclusions

To sum up, the suffix -ATA turns out to be a packaging operator. It selects a rather well-defined set of the verbal lexicon, namely durative and unbounded predicates, and produces derivatives provided with the features of boundedness and point-likeness. Following the terminology introduced in section 1 above, we can represent it as in (47):

- (47) -ATA  
 $A_i = [+ \text{dynamic}, + \text{durative}, - \text{bounded}]$   
 $A_e = [+ \text{dynamic}, - \text{durative}, + \text{bounded}]$

As I noted in (8d) above, this is the most complex case of interaction between morphology and semantics. In fact, this suffix has a particular internal actionality, since it selects a particular set of the verbal lexicon on the basis of actional properties; moreover, it also has a particular external actionality, since it produces derivatives that display a rather different set of actional properties compared to the base-predicates. In this perspective, we can distinguish between true action nouns, in which the operation of transposition does not cause relevant effects on the actional semantics of the base-predicates, and the ATA-nominals, in which we find reversal properties with respect to the base-predicates. This is the reason why ATA-nominals usually do not exert lexical blocking with respect to other nominalizing suffixes. Besides this productive core, there is a set of lexicalised derivatives that do not share the properties established in (47). They behave like true action nouns; given their elevated age and their non-productivity, they apparently seem to be relics of diachronic processes of lexicalization. Finally, we have seen that similar properties hold true for the denominal nouns, which makes them almost identical to the deverbal ones, so that in several cases we are allowed to speak of a double motivation. It remains, however, a matter of further investigation how the diachronic process of affix generalization (cf. Plank 1981, 43ff) to nominal bases took place, which synchronically resulted in a kind of rule telescoping.

## References

- Aronoff, M. 1976. *Word formation in generative grammar*. The MIT Press, Cambridge MA.
- Bartsch, R. 1981. Semantics and syntax of nominalizations. In: Groenendijk, J. – Janssen, T. – Stockhof, M. (eds): *Formal methods in the study of language*, 1–28. Mathematical Centre Tracts 135, Amsterdam.
- Bartsch, R. 1986. On aspectual properties of Dutch and German nominalizations. In: Lo Cascio, V. – Vet, C. (eds): *Temporal structure in sentence and discourse*, 7–39. Foris, Dordrecht.
- Bertinetto, P.M. 1986. *Tempo, aspetto e azione nel verbo italiano*. Accademia della Crusca, Florence.
- Bertinetto, P.M. 1994. Statives, progressives, and habituals: analogies and differences. In: *Linguistics* 32: 391–423.
- Bertinetto, P.M. – Delfitto, D. 1992. Aspect vs. actionality: some reasons for keeping them apart. In: *Eurotyp Working Papers* 6: 1–28.
- Bordelois, I. 1993. Afijación y estructura temática: *-Da* en español. In: Varela, S. (ed.): *La formación de palabras*, 162–79. Taurus, Madrid.
- Brinton, L. 1993. The aktionsart of English deverbal nominalizations. In: Bianchi, P. – Higginbotham, J. – Squartini, M. (eds): *Temporal reference, aspect and actionality*, vol. 1, 27–42. Rosenberg & Sellier, Turin.
- Castelli, M. 1988. La nominalizzazione. In: Renzi, L. (ed.): 333–356.
- Cate, A.P. ten 1985. *Aspektualität und Nominalisierung*. Lang, Frankfurt/Main–Berne–New York.
- Comrie, B. – Thompson, S. 1985. Lexical nominalization. In: Shopen, T. (ed.): *Language typology and syntactic description*. Vol. 3: *Grammatical categories and the lexicon*, 349–98. Cambridge University Press, Cambridge.
- Crocco Galéas, G. – Iacobini, C. 1993. The Italian parasynthetic verbs: a particular kind of circumfix. In: Tonelli, L. – Dressler, W.U. (eds): *Natural morphology. Perspectives for the nineties*, 127–42. Unipress, Padua.
- Dietrich, W. 1994. Grundfragen einer funktionellen Wortbildungslehre, dargestellt am beispiel der -ATA- und -ARIA-Bildungen im Französischen und Spanischen. In: Staib, B. (ed.): *Wortbildungslehre*, 93–111. Münster Logbuch zur Linguistik 5. Münster.
- DISC = *Il Dizionario Italiano Sabatini-Coletti in CD-Rom*. Giunti, Florence, 1997.
- Dixon, R. 1991. *A new approach to English grammar, on semantic principles*. Clarendon, Oxford.
- Ehrich, V. 1991. Nominalisierungen. In: von Stechow, A. – Wunderlich, D. (eds): *Semantik: ein internationales Handbuch der zeitgenössischen Forschung*, 441–58. Walter de Gruyter, Berlin–New York.
- Fleischer, W. – Barz, I. 1992. *Wortbildung der deutschen Gegenwartssprache*. Niemeyer, Tübingen.

- Gaeta, L. 1998. La nominalizzazione deverbale: morfologia e semantica. Ph.D. dissertation, III University of Rome, Rome.
- Gaeta, L. 1999. Un buon argomento contro il separatismo: il suffisso italiano *-anza/-enza*. In: Benincà, P. – Mioni, A.M. – Vanelli, L. (eds): *Fonologia e morfologia dell'italiano e dei dialetti d'Italia. Atti del XXXI Congresso Internazionale di Studi della Società di Linguistica Italiana (Padova, 25–27/9/1997)*, 551–85. Bulzoni, Rome.
- Gatti, T. – Togni, L. 1991. A proposito dell'interpretazione dei derivati in *-ata* e in *s-*. In: *Arbeitspapier Nr. 30, Fachgruppe Sprachwissenschaft. University of Konstanz, Konstanz*.
- Giacalone Ramat, A. 1974. I derivati latini in *-tūra*. In: *Rendiconti dell'Istituto Lombardo* 108 : 236–94.
- Giacalone Ramat, A. 1975. Probleme der lateinischen Wortbildung: Das Suffix *-tūra*. In: Rix, H. (ed.): *Flexion und Wortbildung. Akten der V. Fachtagung der Indogermanischen Gesellschaft*, 120–29. Reichert, Wiesbaden.
- Gross, M. 1981. Les bases empiriques de la notion de prédicat sémantique. In: *Languages* 63 : 7–52.
- Jackendoff, R. 1991. Parts and boundaries. In: *Cognition* 41 : 9–45.
- La Fauci, N. 1979. Costruzioni con verbo operatore in testi italiani antichi. *Esplorazioni sintattiche*. Giardini, Pisa.
- Lüdtke, J. 1978. Prädikative Nominalisierungen mit Suffixen im Französischen, Katalanischen und Spanischen. Niemeyer, Tübingen.
- Marchand, H. 1964. A set of criteria for the establishing of derivational relationship between words. In: *Indogermanische Forschungen* 69 : 10–19.
- Mayo, B. – Schepping, M. – Schwarze, C. – Zaffanella, A. 1995. Semantics in the derivational morphology of Italian: implications for the structure of the lexicon. In: *Linguistics* 33 : 883–938.
- Mourelatos, A. 1978. Events, processes and states. In: *Linguistics and Philosophy* 2 : 415–34.
- Paprotté, W. 1988. A discourse perspective on tense and aspect in standard Modern Greek and English. In: Rudzka-Ostyn, B. (ed.): 447–505.
- Plag, I. 1998. The polysemy of *-ize* derivatives: On the role of semantics in word-formation. In: Booij, G. – van Marle, J. (eds): *Yearbook of Morphology 1997*, 219–42. Kluwer, Dordrecht.
- Plank, F. 1981. Morphologische (Ir-)Regularitäten. Gunter Narr, Tübingen.
- Rainer, F. 1989. I nomi di qualità nell'italiano contemporaneo. Braunmüller, Vienna.
- Rainer, F. 1993. Spanische Wortbildungslehre. Niemeyer, Tübingen.
- Renzi, L. 1988. Grande grammatica italiana di consultazione, vol. I. Il Mulino, Bologna.
- Rohlf's, G. 1969. Grammatica storica della lingua italiana e dei suoi dialetti. Vol. III: Sintassi e formazione delle parole. Einaudi, Turin.



- Rudzka-Ostyn, B. 1988. *Topics in cognitive linguistics*. Benjamins, Amsterdam-Philadelphia.
- Salvi, G. 1988. *La frase semplice*. In: Renzi, L. (ed.): 29-113.
- Samek-Ludovici, V. 1997. A unified analysis of noun- and verb-based italian nominalizations in *-ata*. In: *Arbeitspapier Nr. 80, Fachgruppe Sprachwissenschaft*. University of Konstanz, Konstanz.
- Scalise, S. 1983. *Morfologia lessicale*. CLESP, Padua.
- Simone, R. 1993. *Stabilità e instabilità nei caratteri originali dell'italiano*. In: Sobrero, A. (ed.): *Introduzione all'italiano contemporaneo. Le strutture*, 41-100. Bari, Laterza.
- Smith, C. 1997. *The parameter of aspect (Second edition)*. Kluwer, Dordrecht.
- Symanek, C. 1988. *Categories and categorization in morphology*. Redakcja Wydawnictw Katolickiego Uniwersytetu Lubelskiego, Lublin.
- Talmy, L. 1988. *The relation of grammar and cognition*. In: Rudzka-Ostyn, B. (ed.): 165-205.
- Tekavčić, P. 1972. *Grammatica storica dell'italiano. III volume: Lessico*. Il Mulino, Bologna.
- Thornton, A. 1990. *Vocali tematiche, suffissi zero e 'cani senza coda' nella morfologia dell'italiano contemporaneo*. In: Berretta, M. - Molinelli, P. - Valentini, A. (eds): *Parallela 4*, 43-52. Gunter Narr, Tübingen.
- Ullmer-Ehrich, V. 1977. *Zur Syntax und Semantik von Substantivierungen im Deutschen*. Scriptor, Kronberg/Ts.
- Vendler, Z. 1967. *Linguistics in philosophy*. Cornell University Press, Ithaca NY.
- Wierzbicka, A. 1988. *The semantics of grammar*. Benjamins, Amsterdam-Philadelphia.

Address of the author: Livio Gaeta  
 Dipartimento di Scienze del Linguaggio  
 University of Turin  
 via S. Ottavio 20  
 I-10124 Turin  
 Italy  
 gaeta@cisi.unito.it



## ON APPRECIATIVE SUFFIXES\*

LLUÏSA GRÀCIA AND LÍDIA TURON

### Abstract

In this paper we demonstrate that appreciative suffixes (AS) in Catalan, as well as in other Romance languages, belong to a single group of affixes, independently of the lexical root that they adjoin to. We focus our attention on the semantics of nouns and adjectives that contain an AS. We justify the apparent differences among the derived words by proposing that the AS acts as a modifier of some features of the lexical conceptual structure of the root. We assume Jackendoff's proposal (1990; 1991) for nouns, and we extend the features [ $\pm b$ ,  $\pm i$ ] (bounded and individual) as well as the dimensional feature to adjectives. Only roots containing a [+b] feature will be able to be modified by an AS.

### Introduction

Most of the Romance appreciative suffixes<sup>1</sup> (henceforth AS) can be adjoined to nouns, adjectives and verbs. Although nowadays the productivity of the particular suffixes is different for each category, from a diachronic point of view, it can be said that the Catalan suffixes in (1) have been able to attach to all three categories:<sup>2</sup>

\* We thank our colleagues R. Bayà, J. de Caesaris and O. Fullana for their comments and help. We are also grateful to an anonymous reviewer and to people who attended the 8th International Morphology Meeting (Budapest, June 1998) and the Second Workshop on Lexical Semantics System (Pisa, April 1998), particularly to W. U. Dressler, for their comments. Special thanks to R. Lieber and the English Department of the University of New Hampshire for their support during the three months visit of one of us. This work has been supported by three grants of the Spanish Ministerio de Educación y Cultura (DGICYT PB93-0546-C04-01, DGES PB96-0457-C03-01 and PR95-521) and by three grants of the Direcció General de Recerca of the Generalitat de Catalunya (1995SGR 00113, 1997SGR 00193 and 1997FI 00670).

<sup>1</sup> In this paper we leave aside the prefixes that have appreciative or intensive values, like *mini-*, *hyper-*, etc. In the future, it would be interesting to investigate the differences and similarities between the suffixes and the prefixes with appreciative value. See Bayà (1997) for an approach to this question based on Catalan data.

<sup>2</sup> Lang (1992) and Monterrubio (1990) observe the same phenomenon in Spanish; Scalise (1994) and Napoli-Reynolds (1994) in Italian; and Mascaró (1985) in Catalan.

|      |           |                              |                          |                                                    |
|------|-----------|------------------------------|--------------------------|----------------------------------------------------|
| (1)  | -às       | cotxe                        | cotx + <i>às</i>         | 'big car'                                          |
|      |           | gran                         | grand + <i>às</i>        | 'very big'                                         |
|      |           | allargar                     | allarg + <i>ass</i> + ar | 'to lengthen excessively'                          |
| -isc | astre     | aster + <i>isc</i>           |                          | 'asterisk (small <i>aster</i> )'                   |
|      | moro      | mor + <i>isc</i>             |                          | 'Moslem ('moro') convert to Christianity, Moorish' |
| -on  | ploure    | plov + <i>isc</i> + ar       |                          | 'to rain slightly'                                 |
|      | carrer    | carrer + <i>ó</i>            |                          | 'small street'                                     |
|      | gras      | grass + <i>ó</i>             |                          | 'not very fat'                                     |
| -ic  | enllardar | enllard + <i>on</i> + ar     |                          | 'to make very dirty'                               |
|      | gos       | goss + <i>ic</i>             |                          | 'small dog'                                        |
| -et  | pobre     | pobr + <i>ic</i>             |                          | 'not very poor'                                    |
|      | plorar    | plor + <i>ic</i> + ar        |                          | 'to groan as crying'                               |
|      | llibre    | llibr + <i>et</i>            |                          | 'small book'                                       |
| -atx | alt       | alt + <i>et</i>              |                          | 'not very tall'                                    |
|      | espella   | espell + <i>et</i> + ar      |                          | 'to skin (little pieces)'                          |
|      | poble     | popul + <i>atx</i> + o       |                          | 'populance'                                        |
|      | prim      | prim + <i>atx</i> + <i>ó</i> |                          | '(a little) badly slim'                            |
|      | esprimar  | esprim + <i>atx</i> + ar     |                          | 'to get slim in a bad way'                         |

As can be seen, in all the examples the suffix has a very similar meaning despite the lexical category of the root it adjoins. Traditional grammars usually call the suffixes **diminutive**, **augmentative** and **despective** when adjoining nominal roots, and **intensifiers** in the other cases. Since structurally ASs behave similarly whatever the root is, we agree with the linguists who defend the claim that there is no difference among them. We will not discuss the structural facts here, but will concentrate our attention on the semantics.

The general idea we defend here is that the slightly different interpretations we can find in words containing an AS can easily be explained by a grammar. We assume that all morphemes have a lexical entry in the lexicon in which the semantics is represented by a lexical conceptual structure (henceforth LCS). The lexical process that adjoins a suffix to a root combines the two LCSs. We will see that the AS behaves as an operator that modifies one feature of the LCS of the root, and that this feature is always the same. The various interpretations derive from the various values of the feature. In this paper we will offer an explanation based on nouns and adjectives. Further research should be done in order to include verbs.

### 1. A first approach

As it can be seen in the examples in (1), the AS can modify various aspects of the root: the size of an object, the intensity of a quality, an action or a

process, etc. If we maintain that the affix is the same for all categories, we need to provide an explanation for these different meanings.

From the examples in (2) and (3), we might think that the differences in meaning are only related to the lexical or semantic category of the root:

- |     |     |                                             |                                             |
|-----|-----|---------------------------------------------|---------------------------------------------|
| (2) | (a) | cas (N) + eta (DIM)                         | 'small N ( <i>house</i> )'                  |
|     | (b) | alt (A) + et (DIM)                          | 'not very A ( <i>tall</i> )'                |
|     | (c) | plov (V) + isc (DIM) + ar                   | 'to V ( <i>rain</i> ) slightly'             |
| (3) | (a) | [ <sub>thing</sub> casa ] + [DIM eta ]      | 'small <i>thing</i> ( <i>house</i> )'       |
|     | (b) | [ <sub>property</sub> alt ] + [DIM et ]     | 'not very <i>property</i> ( <i>tall</i> )'  |
|     | (c) | [ <sub>event</sub> plov ] + [DIM isc ] + ar | ' <i>event</i> ( <i>to rain</i> ) slightly' |

Nevertheless, the examples in (4) and (5) show that some other factors are necessary to explain the interpretations of derived words containing an AS:

- |     |                       |                                   |
|-----|-----------------------|-----------------------------------|
| (4) | aigü (N) + eta (DIM)  | '*small water'                    |
|     | mort (A) + et (DIM)   | '*not very dead'                  |
| (5) | marid (N) + às (AUG)  | 'a big husband / a good husband'  |
|     | rector (N) + às (AUG) | 'a big rector / a good rector'    |
|     | cuiner (N) + et (DIM) | 'a small/young cook / a bad cook' |

The data in (4) show that not all nouns or adjectives can be roots for an AS with a denotative meaning. This implies that the suffix is sensitive to some semantic features of the root. On the other hand, (5) shows that some nouns can have more than one interpretation. In one of them, the nouns behave like *house*, and the size is modified (*big husband*), but in the other interpretation what the suffix modifies is some kind of quality related to the noun (*good husband*). It is clear that the meaning of this kind of noun is more complex than the meaning of *house*. We could relate this fact to the verbal origin of *rector* and *cook*, but this explanation does not account for the case of *husband*. We need to accept that the LCSs of *husband*, *rector* and *cook* are more complex than might be thought at first glance. We have to postulate the existence of semantic features in the LCSs that partially determine the meaning of the word. The ASs will be able to modify some of these features. Consequently, derived words will have slightly different values depending on the concrete value of the modified feature and on the semantics of the root. We assume Jackendoff's (1990; 1991) theory, with some modifications.

## 2. Semantic values of ASs

Leaving aside the cases in which the suffix has a connotative value,<sup>3</sup> most denotative meanings of ASs belong to one of these classes: in one case the suffix modifies the extension of what the root denotes, and in the other case it modifies the intensity. Let us analyse each of these two groups separately.

### 2.1. Extension

ASs can modify the extension of the object denoted by the root. According to the semantics of the root, four groups can be distinguished:

- (a) modification of physical extension of the object (in space),
- (b) modification of the number of elements of a set,
- (c) modification of physical extension of the elements of a conglomerate,
- (d) modification of temporal extension.

#### 2.1.1. Physical extension

As for the modification of physical extension, it is obvious that only lexical elements denoting objects with physical extension will allow it. This means that we can find this kind of modification with concrete nouns such as *house*, *book*, *street*, *man*, etc. The derived form means “small/big N”.

According to Jackendoff’s (1991) theory, all these nouns belong to the semantic category **thing** and they have the conceptual features [+bounded, –internal structure] ([+b, –i]). They are [+b] because the entity has bound-

<sup>3</sup> We assume that connotative meanings correspond to “the personal or emotional associations which are suggested by words”, whereas denotation is “the relationship between words and entities in the world to which they refer” (Crystal 1992, 80). Thus, we will consider affective and stylistically marked values of ASs (irony, etc.) as connotatives; the cases in which there is a real diminution or increase will be considered denotative. Consequently, we also consider denotative the cases in which the AS is used as a valorative of a property or a function, as in nouns that designate people. A *marassa* (‘mother+AUG’) refers to a very good (AUG) mother, and an *actoret* (‘actor+DIM’) can refer to a bad (DIM) actor. In both cases, the denotative meaning of the AS is maintained. See Dressler–Merlini Barbaresi (1994) for more details about denotative and connotative meanings of ASs.

aries (as opposed to mass nouns, which are [-b]). The [-i] feature means that the entity is viewed as lacking internal structure (in contrast to nouns like *group*). Moreover, nouns such as *house* or *book* can be analysed in terms of their dimensionality. A line, for example, has dimension-1, a surface, dimension-2, and a ball, dimension-3.

When an AS adjoins a [+b, -i] noun, the suffix modifies the feature [+b], that is, the boundaries of the object. As a consequence, the extension of the object becomes bigger or smaller. If the object has more than one dimension, usually all of them are modified, but it is also possible to modify only some of them: a *llibret* ('small book') may be thick, but in this case its page size has to be small.<sup>4</sup> We summarise these cases in (6):

(6) PHYSICAL EXTENSION

|     |        |           |                    |         |
|-----|--------|-----------|--------------------|---------|
| Ex. | línia  | lini+eta  | 'small/short line' | (dim.1) |
|     | platja | platg+eta | 'small beach'      | (dim.2) |
|     | casa   | cas+eta   | 'small house'      | (dim.3) |

|                             |                         |
|-----------------------------|-------------------------|
| Semantic features:          | [+b (phys.dim.n>0), -i] |
| Lexical category:           | nouns                   |
| Affected semantic features: | [+b (phys.dim. n>0)]    |

Some concrete nouns refer to human entities that have some social function (*president, rector*) or to family relationships (*husband, mother, brother*). These nouns have the features [+b, -i], as they refer to individual physical objects (persons), and then they can have the interpretation in (6) when an AS adjoins them. This interpretation corresponds to one of the two we had in (5). Nevertheless, this is not the usual meaning for them. We will account for other readings later.

### 2.1.2. Number of elements of a set

Some concrete and count nouns have the feature [+b], but, in contrast to the ones we saw before, they are [+i]. The referents are understood as having

<sup>4</sup> Note that when the object has zero dimension, as *punt* 'dot' from a mathematical point of view, an AS cannot modify a dimension that does not exist. A *puntet* would be a little dot, but technically a *punt* 'dot' cannot increase or decrease. Nevertheless, words like *puntet* 'little dot' or *puntàs* 'big dot' are possible. This is because we view dots as having a certain dimension. Jackendoff (1991) uses the notion of "expanded dimensionality" to explain cases similar to this.

internal structure formed by discernable elements. We are referring to nouns such as *group*, *committee*, *bunch*, and so on. The boundaries of the entities denoted by these nouns are determined by the number of elements they contain. As a consequence, modifying the boundaries means to increase or decrease the number of elements. A *grupet* is a group with few elements, a *ramet de flors* ('small bunch of flowers') can be a bunch with few flowers. From that point of view, the AS modifies the feature [+b] as in the previous case.

It is worth noticing that words like *ramet* ('small bunch') can have another reading. It can also refer to a bunch with small flowers. In this case, the AS modifies the [+b] feature contained in the semantic structure of *flower*, the elements that form the set.<sup>5</sup>

We summarise in (7) the interpretations of the derived words in which the AS modifies the number of elements of a set:

(7) NUMBER OF ELEMENTS OF A SET

|     |      |         |               |
|-----|------|---------|---------------|
| Ex. | grup | grup+et | 'small group' |
|     | ram  | ram+et  | 'small bunch' |

|                             |                                  |
|-----------------------------|----------------------------------|
| Semantic features:          | [+b, +i]                         |
| Lexical category:           | nouns                            |
| Affected semantic features: | [+b]                             |
|                             | [+b (phys.dim. n>0)] of the      |
|                             | [+b (phys.dim.n>0), -i] elements |
|                             | that form the set                |

### 2.1.3. Physical extension of the elements of a conglomerate

Some nouns behave in many respects as mass nouns, but they refer to entities formed by perceptible discrete units. We will assign the features [-b, +i] to these **conglomerate** nouns, because, although the objects do not have boundaries, they are understood as having internal structure. Some examples of this kind of nouns are *sorra* ('sand', formed by grains of sand) and *pluja* ('rain', formed by drops of rain).

<sup>5</sup> According to Jackendoff (1991), the LCSs of nouns like *ram* 'bunch', that imply a complement *de* 'of'+ NP, contain a semantic function COMP ('composed of') that the LCSs of the elements of the conglomerate adjoin. This analysis allows him to explain why it is possible to modify a feature of an element that has the function of a complement in the syntax, if it appears.



When an AS is adjoined to a conglomerate noun, it is not possible to modify the boundaries because the noun is [-b]. The AS, as in the case of *ramet* ('small bunch'), modifies the size (the feature [+b]) of the elements that form the conglomerate. For example, *plugeta* (lit. small rain) is a rain with small drops and *sorreta* is a kind of sand with very fine grains.

When a mass noun is not perceived as a conglomerate, that is, when it has the features [-b, -i], the adjunction of an AS cannot modify the size because there is no [+b] feature. The only interpretation for mass nouns like *aigüeta* (lit. small water) or *paperet* (lit. small paper) is the connotative one.<sup>6</sup> For a denotative reading, it is necessary to recategorize the noun as a count noun. Then, they behave like *house*, meaning 'small bottle of water' and 'small piece of paper', respectively.<sup>7</sup>

We summarise in (8) the values of the derived words when the AS affects the size of the elements of a conglomerate:

(8) PHYSICAL EXTENSION OF THE ELEMENTS OF A CONGLOMERATE

|     |       |          |                         |
|-----|-------|----------|-------------------------|
| Ex. | sorra | sorr+eta | 'sand with fine grains' |
|     | pluja | plug+eta | 'rain with small drops' |

|                             |                                                                                               |
|-----------------------------|-----------------------------------------------------------------------------------------------|
| Semantic features:          | [-b, +i]                                                                                      |
| Lexical category:           | nouns                                                                                         |
| Affected semantic features: | [+b (phys.dim. n>0)] of the<br>[+b (phys.dim.n>0), -i] elements<br>that form the conglomerate |

<sup>6</sup> An anonymous reviewer pointed out to us that the Catalan word *aigüeta* could be translated into Russian as *vodka*, which is a diminutive noun from *voda* 'water'. Note that the latter two are mass nouns. We think that the diminutive form has been lexicalised, as its meaning is not easily predictable from its components. We will not analyse lexicalised diminutives here. See Turon (1998) for an explanation of Catalan diminutive lexicalised words.

<sup>7</sup> The semantic process that accounts for the use of these nouns as count nouns is the same one we have explained in footnote 5. Jackendoff (1991) proposes a rule of construal that subordinates the LCS of the mass noun [-b, -i] to an operator COMP (composed of). The output of the process is a count noun.

### 2.1.4. Temporal extension

ASs can modify either the physical extension or the temporal extension of the thing referred by the root. Temporal extension concerns age<sup>8</sup> and duration. ASs—especially diminutive suffixes—can adjoin nominal roots referring to human beings or animals. One of the interpretations of these derived nominals is the one that refers to a young person or animal. See the examples in (9) and (10):

- |      |             |                      |
|------|-------------|----------------------|
| (9)  | german + et | 'young brother'      |
|      | don + eta   | 'young woman'        |
|      | pastor + et | 'young shepherd'     |
|      | cuiner + et | 'young cook'         |
| (10) | balen + ó   | 'baby whale'         |
|      | colom + í   | 'young pigeon'       |
|      | gall + et   | 'young / small cock' |

With some kinds of inanimate nouns, the ASs can modify duration. These nouns denote or imply some kind of event, as we can see in the examples in (11):<sup>9</sup>

- |      |             |                |
|------|-------------|----------------|
| (11) | viatg + et  | 'short trip'   |
|      | curs + et   | 'short course' |
|      | partid + et | 'short match'  |
|      | film + et   | 'short film'   |

Some nouns denote a precise duration. In this case, the literal denotative interpretation is not allowed. See the examples in (12):

- |      |            |                 |
|------|------------|-----------------|
| (12) | hor + eta  | 'short hour'?   |
|      | minut + et | 'short minute'? |
|      | any + et   | 'short year'?   |

<sup>8</sup> According to Dressler and Merlino Barbaresi (1994, 124) the possibility of modifying the age is quite usual in Italian with nouns of family or personal relationships:

- |      |              |                   |
|------|--------------|-------------------|
| (i)  | mamma        | 'young mother'    |
|      | mogliettina  | 'young wife'      |
| (ii) | amichetto/a  | 'young friend'    |
|      | compagnuccio | 'young colleague' |
|      | padroncino/a | 'young boss'      |

<sup>9</sup> As W. U. Dressler pointed out to us (p.c.), some of the examples in (11) can have another reading in Italian and in German: It. *viagg-etto* 'trip+DIM' or G. *Film-chen* 'film+DIM' may be quite long but unimportant ones. Although this reading is possible for some Catalan examples, we think that it is not a usual interpretation for them.

A literal interpretation in (12) is not possible, because the number of elements that form the period cannot be reduced without the noun stopping to be that particular period. As Dressler and Merlini Barbaresi (1994) point out, the only possibility is a reduction in precision. For example, an *horeta* (from *hour*) is a period of approximately 60 minutes, but it can be 50 or 70. The diminution affects the precision of the duration. On the other hand, when the nouns do not refer to a precise duration, the diminutive suffix can shorten the duration, as in (13):

- (13) *eston* + *eta*            'short while'  
       *temporad* + *eta*        'short spell'

The question of which semantic feature is affected is more complicated in these cases. As for animate nouns, we can say that the [+b] feature can refer to two kinds of boundaries: physical ones and temporal ones. Temporal boundaries relate to the fact that animate beings have a limited existence. In this sense, the LCSs of these nouns will have two dimensional values: the ones concerning size and the other ones concerning time. When the AS modifies the [+b] feature of the root, it can modify the physical size as well as the temporal length. It is interesting to notice that an AS can modify both dimensions at the same time, because, pragmatically, young creatures are usually smaller than adults. In the examples in (14) we show the three possibilities:

- (14) *conill* + *et*            only physical dimension  
                                       if 'adult but small rabbit'  
       *german* + *et*            only temporal dimension  
                                       if 'the youngest brother, but very tall'  
       *pastor* + *et*            physical and temporal dimension  
                                       if 'young and small shepherd'

The nouns in (13) and (14) that denote duration do not refer to a material entity and so they do not have physical boundaries. Nevertheless, they have temporal extension with some boundaries. Nouns in (14) are similar to [+b, +i] nouns such as *group* or *bunch*. That is, they are composed of discernible elements. The difference between *hour* and *group* is that the number of elements that form an hour is fixed and the number of elements in a group is not. When an AS adjoins *group*, it affects the [+b] feature in the sense that the number of the elements in the group changes. Since in words like *hour* you cannot modify the number of elements without becoming something other than an hour, the only interpretation is a reduction in precision.

Nouns denoting non-precise temporal extension, such as those in (13), can be the root for an AS as well. In this case, the AS modifies the [+b] feature contained in the semantic structure of the root and the length of the duration can increase or decrease. See in (15):

- (15) una eston + eta 'a short while'  
 una eston + assa 'a long while'

Finally, nouns related to some kind of event, such as those in (11), allow an AS modifying duration. At this point we are not able to offer a detailed analysis, but we do believe that they must be analysed in conjunction with verbal forms, as there is an event implied. Nevertheless, we suggest the possibility of having a [+b] feature related to the telic value at some level of the LCS. The [+b] feature could be included in a secondary LCS present in the LCS of the noun and corresponding to the event. The event would be telic, that is, bounded, as the noun refers to the result of the event.

Summing up, modification of temporal extension parallels modification of physical extension. The AS affects the [+b] feature and, more specifically, the temporal dimension features. We represent this kind of modification in (16):

(16) TEMPORAL EXTENSION

- |     |    |        |             |                          |
|-----|----|--------|-------------|--------------------------|
| Ex. | a. | pastor | pastor + et | 'young shepherd'         |
|     |    | balena | balen + ó   | 'baby whale'             |
|     | b. | estona | eston + eta | 'short while'            |
|     |    | hora   | hor + eta   | 'approximately one hour' |
|     | c. | curs   | curs + et   | 'short course'           |

- Semantic features:
- |    |                                   |
|----|-----------------------------------|
| a. | [+b (phys.dim.n>0, temp.dim), -i] |
| b. | [+b (temp.dim), ±i]               |
| c. | [... +b ...]                      |

- Lexical category: nouns  
 Affected semantic features: [+b (temp.dim.)]

## 2.2. Intensity

The lexical elements that can be modified for intensity are mostly verbs and adjectives. However, as we will see later, some nouns can be modified in this sense as well. The semantic category implied in the modification of intensity is **property** or **situation** (*states* and *events*). In this paper we will leave aside

verbs. We begin analysing derived adjectives and then intensity in nouns will be explained.

### 2.2.1. The intensity of a property

Lexical items denoting a property are mostly adjectives. Jackendoff's work does not, however, develop a system of features for adjectives. As we would like to present a homogeneous explanation for our data, we will try to extend Jackendoff's system to adjectives.

One of the traditional classifications of adjectives distinguishes gradable from nongradable adjectives. Gradable adjectives, such as *tall*, *sad* or *nice*, can be modified for grade: one can say *very tall*, *not very nice*, and so on. On the other hand, *dead* and *Cuban* are nongradable adjectives, and it is not possible to say *very dead* or *not very Cuban* unless a metaphorical sense is used.<sup>10</sup>

Using the features we have seen for nouns, we can distinguish these two kinds of adjectives by means of the  $[\pm b]$  feature and the associated dimensionality feature. A  $[+b]$  property is a property with some extension and boundaries. In other words, the property *tall*, for example, is a property that extends between two boundaries; there are several points for which you can say that an entity has this property. At the different points the property is *more* or *less* intensive, depending on which boundary it is near. We can consider that *tall* is a property with a 1-dimension.

Nongradable adjectives, on the other hand, can be seen as properties with boundaries but with a 0-dimension. Since the dimension is zero, the property cannot be placed at varying points and boundaries cannot change.

When an AS adjoins a gradable adjective root with a  $[+b]$  feature and a 1-dimension, the semantic effect is that the property is located near the higher boundary of the scale (if augmentative) or near the lower one (if diminutive). The diminutive form for *tall* means 'not very tall', that is, the property is located at a low point on the scale. In the augmentative form, meaning 'very tall', the property is located at a high point on the scale. As in the cases analysed before, the affected feature is  $[+b]$ .

<sup>10</sup> It is possible to say *molt mort* lit.: 'very dead' meaning 'very tired'. As is well known, relational adjectives such as *Cuban*, *Spanish*, *Italian* cannot be modified for grade in their normal use. Nevertheless, they accept grade modifiers when used as non-relational adjectives: *un home molt espanyol* 'a very Spanish man' is a man who has many properties typically attributed to Spanish people.

As predicted, nongradable adjectives cannot be the base for an AS if they are not used as gradable adjectives.<sup>11</sup> Sequences such as those in (17) are not well-formed:

- (17) \*arbre mort + et            'dead + DIM tree'  
       \*passaport cuban + et    'Cuban + DIM passport'

Although the adjective is [+b], these forms are impossible because the dimension is zero and then it is not possible to change the boundaries if no rule can increase the dimension.

In (18) we summarise the values of derived words corresponding to a property in which an AS modifies its intensity:

(18) INTENSITY OF A PROPERTY (ADJECTIVE ROOTS)

- Ex. alt    alt + et            'not very tall' (less tall than *tall*)<sup>12</sup>  
       gran    grand + às           'very big' (bigger than *big*)

Semantic features:            [+b (dim.1)]

Lexical category:            adjectives

Affected semantic features: [+b (dim.1)]

### 2.2.2. Intensity in some nouns

At the beginning of the paper (see the examples in (5)), we referred to some nouns, like *husband*, *rector*, or *cousin*, that have different readings when an AS adjoins them. We have already accounted for two of these interpretations:

<sup>11</sup> It is possible to adjoin an AS to a relational adjective when it is used as a gradable qualifying adjective. A phrase like *polític espanyolàs* can refer to a politician who has a very strong Spanish behaviour or world view. And, as an anonymous reviewer pointed out to us, words such as *romanina* ('Roman+DIM') and *Toscanini* ('Tuscany+DIM+plural') exist in Italian. We think, however, that these Italian diminutives are lexicalised forms, because their meanings are not predictable: a *romanina* is a kind of bread (in some areas), and *Toscanini* is a surname.

<sup>12</sup> In the case of the adjectives that indicate a reduced size, as *curt* 'short', the adjunction of an AS with diminutive value causes the selection of the low boundary of the scale and, consequently, the real dimension is bigger. That is, a *camí curtet* 'short+DIM way' is longer than a *camí curt* 'short way'. That proves that diminutives do not necessarily denote smaller dimensions with respect to the base, but rather that they select the low values of a scale of gradation. The adjective *petitet* 'little+DIM' is different. If a book is *petitet* it is smaller than a *petit* book. According to Dressler-Merlini Barbaresi (1994), this is because the denotative meaning of the base coincides with the meaning of the suffix and then it is perceived as an intensification.

modification of physical extension (a *maridàs* can be a 'big husband'), and modification of temporal extension (a *germanet* can be a 'young brother'). Nevertheless, some of these nouns can also be interpreted in yet another way: a *maridàs* can refer to a 'good husband'. Under this interpretation, the size and the age of the person is not modified; rather, what is modified is the quality of his "function" as a husband.

As these are nouns that refer to human beings (physical objects), they have the features [+b, -i] and the AS can modify the [+b] feature, either the physical dimension or the temporal one. How can we account for the third reading? In this interpretation, nouns like *husband* are similar to adjectives:<sup>13</sup> as they refer to a property (a "function") of a person. Our proposal is that the LCS of nouns like *husband* contains a secondary structure equivalent to the LCS of an adjective modifying the physical object represented by the primary structure. See the example in (19):

$$(19) \left[ \begin{array}{c} \left[ \begin{array}{cc} & +b, -i \\ \text{thing} & \text{MAN} \end{array} \right] \\ \text{BE} \left( \left[ \begin{array}{cc} & +b \text{ dim.1} \\ \text{property} & \text{HUSBAND} \end{array} \right] \right) \\ \text{thing} \end{array} \right]$$

The internal **thing** constituent defines the entity as a physical object (a person, a man). A secondary structure corresponding to a property has been subordinated to the primary structure. The **property** constituent is the argument of a BE function. As we can see, the secondary structure corresponds to that of a gradable adjective. In this way, the interpretation 'good husband' comes from the modification of the property *husband*. The quality (or intensity) of the function of being a husband is affected by the AS. When the AS is augmentative, the intensification is done by selecting the area placed near the positive boundary (*good husband*). Although family nouns do not often allow the corresponding diminutive interpretation, social function nouns do: a *rectoret* can be a nice person who doesn't play the function of being a rector very well.<sup>14</sup>

<sup>13</sup> See the "degree nouns" in Dressler-Merlini Barbaresi (1994, 416).

<sup>14</sup> The interpretation with negative value is more limited than the one with positive value. It is difficult to interpret nouns like *mareta* 'mother+DIM', *maridet* 'husband+DIM', *germanet* 'brother+DIM', etc., with the value of bad mother, bad husband or bad brother. Probably the explanation of this fact would be pragmatic: the negative value is incompatible with the affective value that these diminutives related to close family relationships often carry.

Finally, in (20) we present another kind of nouns that allow an AS modifying intensity:

- (20) soroll + et 'weak noise'  
 calor + assa 'strong heat'  
 olor + eta 'weak smell'  
 vent + às 'strong wind'  
 llumen + eta 'weak light'

The nouns in (20) refer to physical entities that you cannot touch but that your senses can perceive. Their "extension" is their intensity: a 'big wind' is a 'strong wind', a 'small noise' is a 'weak noise'. As in the previous cases we suppose that the LCS of the nouns in (20) contains a [+b] feature that accounts for this gradation.

In (21) we summarise the modification of intensity in nominal roots:

(21) INTENSITY OF A PROPERTY (ADJECTIVE ROOTS)

- Ex. a. marit marid + às 'good husband'  
 b. vent vent + às 'strong wind'

- Semantic features a.  $\left[ \begin{array}{l} +b, -i \\ [+b \text{ (dim.1)}] \end{array} \right]$   
 b.  $[\dots +b \text{ (dim.1)} \dots]$

Lexical category: nouns  
 Affected semantic features: [+b (dim.1)]

### 3. The LCS of AS and the derivational process

As we have seen, ASs adjoin nominal and adjectival roots (and, we assume, also verbal roots) and modify a [+b] feature with a dimension superior to zero. The different nuances in the interpretation depend on the concrete value of the [+b] feature, that is, on the kind of boundaries it refers to (physical or temporal extension, scale of gradability, etc.). The LCSs we propose for the affixes are represented in (22):

- (22) LCS of a diminutive suffix      LCS of an augmentative suffix  
 [ DIM ( [  $\alpha$  +b (dim.n>0) ] ) ]    [ AUG ( [  $\alpha$  +b (dim.n>0) ] ) ]

The functions DIM and AUG can be seen as operators. They have scope over a [+b] feature with a dimensional value higher than zero contained in the



LCS of the root the suffix adjoins. We use the variable ‘alpha’ to indicate any conceptual category (**thing**, **property**, and maybe **event**). That means that ASs can only combine with roots containing a [+b] feature with a dimension higher than zero in their LCSs. This allows us to explain all the derived words we have analysed as well as the impossibility of having a denotative interpretation with other affixed words. [-b] nouns like *water* or [+b, dim.0] roots like *point*, *instant* or *Cuban* do not allow the denotative reading of the AS.

Let us present some examples that show how the LCSs of the root and the affix combine. See (23) and (24):

- (23) casa     [thing +b (phys.dim.3), -i]                    ‘house’  
 caseta     [thing DIM (+b (phys.dim.3)), -i]            ‘small house’

- (24) actor ‘actor’

$$\left[ \begin{array}{l} \left[ \begin{array}{l} \text{thing} \\ \text{MAN} \end{array} \right] \text{ +b (phys.dim.3, temp.dim.1), -i} \\ \text{BE} \left( \left[ \begin{array}{l} \text{property} \\ \text{ACTOR} \end{array} \right] \text{ +b (dim.1)} \right) \\ \text{thing} \end{array} \right]$$

- a. actoret ‘small person who is an actor (Woody Allen)’

$$\left[ \begin{array}{l} \left[ \begin{array}{l} \text{thing} \\ \text{MAN} \end{array} \right] \text{ +b (DIM (phys.dim.3), temp.dim.1), -i} \\ \text{BE} \left( \left[ \begin{array}{l} \text{property} \\ \text{ACTOR} \end{array} \right] \text{ +b (dim.1)} \right) \\ \text{thing} \end{array} \right]$$

- b. actoret ‘young person who is an actor (Macaulay Culkin)’

$$\left[ \begin{array}{l} \left[ \begin{array}{l} \text{thing} \\ \text{MAN} \end{array} \right] \text{ +b (phys.dim.3), DIM (temp.dim.1), -i} \\ \text{BE} \left( \left[ \begin{array}{l} \text{property} \\ \text{ACTOR} \end{array} \right] \text{ +b (dim.1)} \right) \\ \text{thing} \end{array} \right]$$

- c. actoret 'person who is a bad actor (a bad amateur actor)'

$$\left[ \begin{array}{c} \left[ \begin{array}{c} +b (\text{phys.dim.3, temp.dim.1), -i \\ \text{thing MAN} \end{array} \right] \\ \text{BE} \left( \left[ \begin{array}{c} +[\text{DIM} (+b (\text{dim.1}))] \\ \text{property ACTOR} \end{array} \right] \right) \end{array} \right]_{\text{thing}}$$

In (23) the AS modifies the [+b (phys.dim.3)] feature of the root. As it is a physical dimension, the physical extension of the object is modified in one, two or three dimensions.

The three possible interpretations in (24) are explained by the fact that the AS can combine with three different features. In (24a) it combines with the [+b (phys.dim.3)] primary feature and the meaning is "small man who is an actor". In (24b) the affected feature is [+b (temp.dim.1)], and the derived word refers to a young man who is an actor. Finally, in (24c) the AS modifies the [+b (dim.1)] feature of the subordinate property and the word refers to a man who is a bad actor.

#### 4. Conclusions

The aim of this paper was to demonstrate that ASs that adjoin nominal roots are the same that adjoin adjectival (and probably verbal) roots. We have only analysed them from a semantic point of view. Following the general idea in Jackendoff (1990, 1991), we concluded that ASs behave as operators that modify a [+b (dim.n>0)] feature present in the LCS of the root. The different nuances in the semantic interpretation of derived words come from the kinds of boundaries the [+b] feature refers to. The boundaries can be physical, temporal or gradual. The way we started here should be completed with a semantic analysis of verbs containing an AS. Our prediction is that ASs should behave as they do with all other lexical categories.

### References

- Bayà, M. 1997. La prefixació en català: aproximació teoricodescriptiva. Master's thesis, Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra, Barcelona.
- Crystal, D. 1992. An encyclopedic dictionary of language and languages. Blackwell, Cambridge MA.
- Dressler, W. – Merlini Barbaresi, L. 1994. Morphopragmatics, diminutives and intensifiers in Italian, German, and other languages. Mouton de Gruyter, Berlin.
- Jackendoff, R. 1990. Semantic structures. The MIT Press, Cambridge MA.
- Jackendoff, R. 1991. Parts and boundaries. In: Levin, B. – Pinker, S. (eds): Lexical and conceptual semantics, 9–45. Blackwell, Cambridge MA.
- Lang, M. 1992. Formación de palabras en español. Cátedra, Madrid.
- Mascaró, J. 1985. Morfologia. Enciclopèdia Catalana, Barcelona.
- Monterrubio, J. 1990. La disminución léxica: un concepto exclusivamente nominal? In: Verba 17: 77–92.
- Napoli, D. – Reynolds, B. 1994. Evaluative affixes in Italian. In: Yearbook of Morphology 1994, 151–78. Kluwer, Dordrecht.
- Scalise, S. 1994. Morfologia. Il Mulino, Bologna.
- Turon, L. 1998. Els diminutius lexicalitzats en l'àmbit lexicogràfic. Ms. Universitat de Girona / Universitat Pompeu Fabra, Girona / Barcelona.

Address of the authors: Lluïsa Gràcia and Lídia Turon  
Universitat de Girona  
Departament de Filologia i Filosofia  
Pl. Ferrater Mora, 1.  
17071 Girona  
gracia@skywalker.udg.es  
turon@skywalker.udg.es



## WHERE PLURAL REFUSES TO AGREE: FEATURE UNIFICATION AND MORPHOLOGICAL ECONOMY

ALBERT ORTMANN

### Abstract

The paper offers an account of languages such as Hungarian which do not allow for number concord, that is, for several noun-phrase internal realisations of plural. It is looked at the combinations of noun and adjective, numeral and noun, and subject and verb. I first show that an underspecification analysis fails to capture the data. I argue that nevertheless, the lack of number agreement in this language type is only apparent and that it is still possible to stick to the concept of feature unification for these constructions. The solution I propose is coached in the framework of Optimality Theory and crucially relies on an economy constraint which I call PEPL and which outranks two other constraints: MAP, which requires a correspondence of semantic aggregate individuation and the morphological feature [+pl]; and REALISE( $\mu$ ), which requires that affix material that fits into the morphosyntactic context should be realised. A consequence of the analysis is that non-default mapping of aggregate semantics to the morphosyntactic specification [-pl] is a typological option in order to respect formal agreement. The variation between 'Type Hungarian' languages and 'Type English' languages (i.e., languages that exhibit plural concord) is thus accounted for in terms of a different ranking of the constraints that require morphological economy (PEPL) and explicitness (MAP, REALISE( $\mu$ )), respectively.

### Introduction<sup>1</sup>

In this paper I provide an analysis of plural constructions in languages such as Hungarian which do not exhibit redundant plural marking of the kind found

<sup>1</sup> This work was carried out in the Special Research Programme SFB 282 "Theorie des Lexikons", financed by the German Science Foundation (DFG). I would like to thank the following for discussion, comments on earlier versions, or help with the data: Kerstin Blume, Ray Fabri, Birgit Gerlach, Martin Haspelmath, Sandra Joppen, Ingrid Kaufmann, Martin Krämer, Antje Schmidt, and Dieter Wunderlich. The Hungarian data are due to Tünde Vallyon, whose cooperation is also gratefully acknowledged. Parts of this work were presented at the 8th International Morphology Meeting, Budapest, June 11-14, 1998, and at "Lexicon in focus: International Conference of the SFB 282 'Theory of the Lexicon'", Wuppertal, August 17-19, 1998. I am grateful to the audiences of both occasions, especially to Jim Blevins, Casper de Groot, and W.U. Gustav Wurzel. Finally, thanks for thorough and constructive criticism are due to an anonymous ALH-reviewer.

in such languages as English. The contexts that will be looked at chiefly are combinations of noun and adjective, numeral and noun, and subject and verb. Contrary to what one might prefer for reasons of restrictiveness, the facts call for a solution that allows for violable constraints with a language-specific ranking. The analysis I propose makes use of an economy constraint which allows one to adhere throughout to the formal means of feature unification.

The paper is organised as follows: section 1 gives an overview of the data and contrasts the multiple realisation of plural in languages with DP-internal number agreement to the behaviour of languages with only one plural marking. Section 2 points out the theoretical relevance of the data and shows why there is no simple explanation for the lack of multiple plural marking in the latter type. In section 3 I sketch two apparently promising analyses and show why these do not adequately account for the facts. Eventually, in section 4, the bulk of the paper, I argue for a solution that draws on violable constraints with a language-specific ranking in the framework of Optimality Theory. The major results are summarised in section 5.

## 1. Two types of DP-internal agreement

When looking at noun-phrase internal agreement in contexts of plural reference, one finds two major strategies in languages with overt number inflection. In this paper, I will refer to these types informally as ‘Type English’ and ‘Type Hungarian’, respectively, according to a prominent representative of each type.

### 1.1. ‘Type English’: several overt realisations of plurality

The first type involves several overt realisations of plural in the noun phrase. In English, for example, a numeral combines with a plural noun and a plural demonstrative. Similarly, the example from German involves three overt realisations of plural next to the numeral. Also, the Bantu concord prefixes, here illustrated from Swahili, involve number information in each token, which is fused with noun class information.

- |     |          |        |                  |       |        |       |                         |          |
|-----|----------|--------|------------------|-------|--------|-------|-------------------------|----------|
| (1) | English: | this   | island           |       | these  | two   | island-s                |          |
|     |          | dem.sg |                  |       | dem.pl | num   | -pl                     |          |
|     | German:  | das    | groß-e           | Haus  | die    | drei  | große-en                | Häus-er  |
|     |          | df.sg  | huge-sg          | house | df.pl  | three | huge-pl                 | house-pl |
|     |          |        | ‘the huge house’ |       |        |       | ‘the three huge houses’ |          |

|          |                 |         |                |         |                 |
|----------|-----------------|---------|----------------|---------|-----------------|
| Swahili: | ki-siwai        | ki-zuri | vi-siwai       | vi-zuri | (Carstens 1991) |
|          | 7-island        | 7-good  | 8-island       | 8-good  |                 |
|          | 'a nice island' |         | 'nice islands' |         |                 |

To account for morphological number agreement in these languages, we can simply assume unification of inflectional features, including number, in terms of compatible feature specifications, as defined for example in HPSG (Pollard–Sag 1994, chapter 2) or in Wunderlich (1994).

## 1.2. 'Type Hungarian': only one overt realisation of plurality

In contrast to the language type just presented, in many other languages, here referred to as 'Type Hungarian', there is no morphological agreement with a plural noun within the noun phrase.

### 1.2.1. No number agreement with plural nouns: adjectives

In Hungarian itself, for example, as in English or German, there is a contrast between an unmarked singular and a marked plural form of the noun; see (2a,b). However, nominal modifiers such as the attributive adjective, see (2c,d), do not exhibit number agreement although the adjective can in principle have a plural specification, as is clear from its predicative use in (2e).

(2) Hungarian:

|          |                      |               |       |      |           |         |
|----------|----------------------|---------------|-------|------|-----------|---------|
| (a) hajó | (b) hajó-k           | (c) egy       | gyors | hajó | (d) gyors | hajó-k  |
| ship     | ship-pl              | a             | fast  | ship | fast      | ship-pl |
| 'ship'   | 'ships'              | 'a fast ship' |       |      | 'ships'   |         |
| (e) A    | hajó-k               | gyors-ak.     |       |      |           |         |
| df       | ship-pl              | fast-pl       |       |      |           |         |
|          | 'The ships are fast' |               |       |      |           |         |

Similarly, in Basque, there is an unmarked singular *etxe-a* and a marked plural *etxe-ak*; see (3a,b). (3c) shows that in adjectival modification contexts only the adjective, but not the noun, exhibits the plural marker. As (3d) shows, it is always the last element in the noun phrase that is marked for plural, so there is no DP-internal plural agreement—although (3e) shows that in principle both the noun and the adjective can have a plural specification, namely in predicative use.

(3) Basque:<sup>2</sup>

- |            |              |                   |             |
|------------|--------------|-------------------|-------------|
| (a) etxe-a | (b) etxe-ak  | (c) etxe          | handi-ak    |
| house-det  | house-det.pl | house             | huge.det.pl |
| 'house'    | 'houses'     | 'the huge houses' |             |
- 
- |                                 |       |       |                  |
|---------------------------------|-------|-------|------------------|
| (d) etxe                        | gorri | handi | ederr-ak         |
| house                           | red   | huge  | beautiful-det.pl |
| 'the red huge beautiful houses' |       |       |                  |
- 
- |                                        |     |             |          |         |
|----------------------------------------|-----|-------------|----------|---------|
| (e) Etxe-a-ren                         | bi  | alde-ak     | txuri-ak | dira.   |
| house-det-gen                          | two | side-det-pl | white-pl | aux.3pl |
| 'The two sides of the house are white' |     |             |          |         |

As for Basque, the analysis of the plural construction is quite obvious: the plural marker is simply a phrasal suffix, attached to the whole noun phrase rather than to the noun. Given this alone, however, the lack of plural agreement in 'Type Hungarian' languages, including Basque, is not sufficiently accounted for, as will be shown in section 2.1.

### 1.2.2. No plural marker at all if 'plurality' is realised somewhere else

In addition to the facts introduced so far, in 'Type Hungarian' languages there is no plural marker at all (that is, not even on the noun itself) if the concept of 'plurality' is signalled by a different category. This is basically the case in two environments, namely nouns combined with numerals or quantifiers, and DP-external agreement. Both environments will be exemplified in turn.

#### 1.2.2.1. Numerals and quantifiers

In Hungarian, the combination of a numeral with a noun yields *öt hajó*, literally 'five ship', instead of \**öt hajó-k*:

- |                |          |              |              |
|----------------|----------|--------------|--------------|
| (4) Hungarian: | egy hajó | öt hajó      | sok hajó     |
|                | one ship | five ship    | many ship    |
|                | 'a ship' | 'five ships' | 'many ships' |

The same holds for many other languages, such as Basque, Kurdish, Persian (the latter two from the Indo-Iranian branch of Indo-European), Turkish, the South Caucasian (Kartvelian) languages and the unrelated North-East Caucasian languages, or Quechua, some of which are illustrated here:

<sup>2</sup> I am grateful to Sandra Joppen for making available to me her informant work on Basque (precisely speaking, the dialect of Gipuskua).



- (5) Basque:            etxe bat            etxe-ak            hiru etxe            \*hiru etxe-ak  
                           house one            house-det.pl        three house        three house-det.pl  
                           ‘one house’        ‘houses’            ‘three houses’
- (6) Kurdish:            mamoste            du mamoste        çar hesp            (Barnas–Salzer 1994)  
                           teacher            two teacher        four horse  
                           ‘a teacher’        ‘two teachers’    ‘four horses’
- (7) Turkish:            yıl                    yıl-lar            sekiz yıl            (Lewis 1967)  
                           year                    year-pl            eight year  
                           ‘year’                    ‘years’            ‘eight years’
- (8) Tsova-Tush (Nakh < NE Caucasian; Holisky–Gagua 1994, 189):
- |           |         |            |              |
|-----------|---------|------------|--------------|
| cha k’nat | k’nat-i | si k’nat   | qo k’nat     |
| one boy   | boy-pl  | two boy    | three boy    |
| ‘one boy’ | ‘boys’  | ‘two boys’ | ‘three boys’ |
- (9) Quechua (Huanca dialect; Cerron-Palomino 1976, 125f):
- |       |            |             |              |
|-------|------------|-------------|--------------|
| mishi | mishi-kuna | tawa mishi  | achka uwish  |
| cat   | cat-pl     | four cat    | many sheep   |
| ‘cat’ | ‘cats’     | ‘four cats’ | ‘many sheep’ |

The lack of plural agreement is particularly striking in Archi, a Dagestan language of the North-East Caucasus. Nominal modifiers do show overt number agreement in Archi, but in the singular rather than in the plural, due to the presence of a numeral. In (10a) and (10b), the singular suffix for class I (masculine) and class II (feminine), respectively, is found on the demonstrative, on the relative clause, on the possessive pronoun, on the numeral itself, and on the adjective.

- (10) Archi: (Kibrik 1994)
- (a) yo-w            zon            L’annu-w    w-is            q’°le-w-u    dol:zu-w    usdu  
       dem-Isg    pron1sg    love-Isg    Isg-pron.gen    two-Isg    elder-Isg    brother  
       ‘these two elder brothers of mine who love me’
- (b) ya-r            zon            L’annu-r    d-is            q’°le-r-u    dol:zu-r    doşdur  
       dem-IIsg    pron1sg    love-IIsg    IIsg-pron.gen    two-IIsg    elder-IIsg    sister  
       ‘these two elder sisters of mine who love me’

A note on the various languages that belong to ‘Type Hungarian’ is in order. These languages are genetically unrelated, though most of them belong to a contiguous geographical area (that is South-East Europe and southern Asia). However, that there is more about the property of lacking plural agreement than simply being a chance areal feature is evidenced by the fact that also

geographically remote languages such as Basque or (varieties of) Quechua exhibit the same pattern. We therefore deal with one option provided by UG next to other options such as that of 'Type English', or of number not being an inflectional category at all (as for example in Vietnamese).

### 1.2.2.2. DP-external agreement

In addition to the facts of noun phrase internal agreement, in subject-verb agreement we find that in Kurdish, in a context of plural reference only the verb, but not the subject, is marked for plural. Thus, the plural suffix of nouns, *-an*, cannot be realised in (11).

- (11) Kurdish: (Akrawy 1982)
- |                                                                                                         |                                                                                                      |
|---------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------|
| <p>(a) Mirov hat.<br/>           man come.past.sg<br/>           'The man came.'</p>                    | <p>(b) Mirov(*-an) hat-in.<br/>           man(*-pl) come.past-3pl<br/>           'The men came.'</p> |
| <p>(c) E zarok dixwîn-in.<br/>           dem child learn-3pl<br/>           'These children learn.'</p> |                                                                                                      |

In virtually all other 'Type Hungarian' languages it is the verb that is incapable of showing plural agreement in a certain plurality context, namely if the subject contains a numeral or quantifier.<sup>3</sup> In other words, in this constellation the singular form is found on the verb, hence *iszik* rather than plural *isznak* in (12), and *goravs* rather than *goraven* in (13b):

- (12) Hungarian: (Az) öt nagynéni sört isz-ik / \*isz-nak.  
 (df) five aunt beer drink-3sg / drink-3pl  
 '(The) five aunts are drinking beer.'
- (13) Georgian: (a) Knut-eb-i gorav-en. (b) Sami knut-i gorav-s.  
                   kitten-pl-nom roll-3pl           three kitten-nom roll-3sg  
                   'The kittens are rolling.'           'Three kittens are rolling.'  
(Harris 1981)
- (14) Tsova-Tush: Si k'nat v-eʔeʔ.  
                   two boy 3sg-came  
                   'Two boys came.' (Holisky-Gagua 1994)

We can conclude that as far as DP-external agreement is concerned, there are two different subtypes of plural being realised on one constituent only:

<sup>3</sup> For the exception of Basque see section 4.5.

- (I) Subject-verb agreement is governed by the form of the lexical head of the subject noun phrase (Hungarian, Georgian, and most others).  
 (II) The specification of the lexical head of the subject noun phrase is governed by the verb; that is, the subject in combination with a plural verb shows singular (Kurdish).

Let me also summarise the observations from section 1.1. In none of the 'Type Hungarian' languages does the adjective exhibit a plural specification, for one of the following reasons:

- (III) The adjective does not inflect at all (Persian<sup>4</sup>), or only with respect to case (Georgian<sup>5</sup>). For these cases, there is no need to account for the lack of number agreement, since there are simply no agreement markers available.  
 (IV) Only the final element of the noun phrase is marked for number (Basque). As a phrase final marker by its very nature is not an agreement marker, it can never appear with non-DP-final adjectives (see the representation in the following section). Hence it is only for nouns preceded by a numeral that the lack of this suffix has to be accounted for.

<sup>4</sup> The invariance of the Persian adjective is illustrated in (i). 'ez' stands for the *ezafe*-suffix, which licenses a noun modifier such as a possessor or an adjective (Amin-Madani-Lutz 1972, 57):

- (i) (a) gol-e           zard           (b) gol-hay-e       zard  
       flower-ez   yellow        flower-pl-ez   yellow  
       'yellow flower'       'yellow flowers'  
       (c) Xub-an           hamise   xosbaxt   hast-and.  
       good-pl.anim   always   happy   cop-3pl  
       'The good are always happy.'

The adjective is identical for singular and plural; cf. (ia) and (ib). As (ic) shows, this also holds for the predicative use.

<sup>5</sup> In contemporary Georgian, adjective agreement with the noun concerns only case, not number (Aronson 1991, 236f):

- (i) (a) ʒvel-i       c'ign-i           (b) ʒvel-i       c'ign-eb-i       (c) ʒvel-ma   c'ign-ma  
       old-nom   book-nom       old-nom   book-pl-nom   old-erg   book-erg  
       'old book'                   'old books'                   'old book'  
       (d) ʒvel-ma   c'ign-eb-ma  
       old-erg   book-pl-erg  
       'old books'

In its predicative use the adjective usually is also uninflected for number, hence similarly to Persian, there is no number morphology to account for.

- (V) The adjective is marked for number only in its predicative use (Hungarian).<sup>6</sup> This calls for an explanation, since here an available agreement marker must not be realised.

A theoretical analysis of the data in (2) to (14) has to account for (I), (II) and (V). In the following section I show that in these cases plural agreement seems indeed to be ‘blocked’ by the grammar, since unlike in the case of (III) and (IV), the morphological potential for agreement is available. Also, the claim made in (IV) concerning the Basque plural suffix will be given more substance.

## 2. The theoretical problem: formal agreement or not for ‘Type Hungarian’?

In order to state the theoretical relevance of the data presented so far, it is essential to see that the problem cannot be reduced to either the status of plural markers as phrasal suffixes, or to a special semantics of the plural markers for ‘Type Hungarian’ languages. Both claims will be justified in turn.

### 2.1. Do ‘Type Hungarian’ plural markers select noun phrases rather than nouns?

From examples (3d,e) above we saw already that there is one language of ‘Type Hungarian’, namely Basque, where the suffix marker always appears suffixed to the whole noun phrase rather than to the noun itself. The Basque plural marker is therefore to be analysed as selecting a phrase rather than a word stem. This is expressed by the following representation.<sup>7</sup>

<sup>6</sup> Strictly speaking, this statement requires some precision: in the absence of a head noun, the attributive adjective is marked for number instead, as well as for all other inflectional categories. The same holds for other ‘Type Hungarian’ languages. The point that is made here is thus even clearer: the adjective is marked for number only if there is no DP-internal agreement. As it seems, the most straightforward account for this distribution would be that adjectives in noun-less constructions are substantivised, that is, they occupy the position of the head noun.

<sup>7</sup> In the notational format of this representation I follow Zimmermann (1998).

- (15) Representation of the Basque definite plural marker *-ak*: /ak/, [PrWd —]  
 ⟨+def,+pl⟩  
 +N,-V,+MAX  
 $\lambda P \lambda x^{(+pl)} [P(x) \ \& \ AGGR(x)]$

As for the phonology, *-ak* is integrated into a prosodic word (PrWd) in the same way as enclitics are. Morphologically, it adds the specification ⟨+def,+pl⟩. However, instead of selecting a noun it selects a maximal projection, namely a noun phrase (+N,-V,+MAX). In other words, it is only phonologically bound to a stem, not morphologically. Semantically, *-ak* can be conceived of as operating not simply on the noun predicate but rather on the more complex predicate expressed by the entire NP; for example, the noun combined with a modifier or a possessor. This characterisation captures the fact that there is no plural agreement in the Basque noun phrase.

Given this representation for Basque, the question is if (15) is a pattern that generally accounts for the lack of plural agreement in 'Type Hungarian' languages. I will briefly point out that an analysis of the plural marker selecting a noun phrase does not carry over to all the data in (1)–(14).

First, for all other languages of this type the position of the plural marker (and also the case marker) is clearly not governed by linearity, hence it does not appear phrase-finally but rather on the head of NP, that is on the noun. This is evidenced by cases where the noun is followed by a modifier such as a relative clause, as in the following example from Hungarian:

- (16) Lát-t-am [DP a gyors, angol hajó-k-at [CP amely-ek tegnap  
 see-pret-1sg df fast English ship-pl-acc rel-pl yesterday  
 Hongkong-ba hatalmas rakomány-ai-k-kal be-fut-ott-ak ]].  
 Hongkong-illative huge good-pl-por3pl-inst arrive-pret-3pl

'I saw the fast English ships that arrived in Hongkong with huge goods yesterday.'

(16) clearly shows that the plural suffix is realised on the head noun, not phrase-finally. The explanation of the lack of plural agreement in Basque is therefore not available for Hungarian and for the other languages of that type.

Second, even for Basque, the selection of a noun phrase rather than a noun stem does not account for one of our crucial problems, namely the lack of the marker in the presence of a numeral in, for example, *hiru etxe* 'three houses'. Nothing would exclude the plural marker, thus *\*hiru etxe-ak* would be expected instead. Therefore, the lack of number agreement in Basque is

by no means captured by the status of the plural marker and calls for a more elaborate solution.

In sum, an analysis of the plural marker selecting a noun phrase rather than a noun stem is not only inappropriate for the other 'Type Hungarian' languages, but it does not even account for all Basque data.

## 2.2. Does 'Type Hungarian' plural differ from 'Type English' in its semantics?

From a cross-linguistic point of view it must be stated that there is quite some variation as to whether or not, to what extent, and for which function the category of number is grammaticalised. For example, as Mithun (1988) points out in much detail, in many North-American languages, such as Central Pomo, the realisation of a number affix on a noun has a distributive function with respect to the event denoted by the verb, rather than simply indicating multiple individuals (in which case the affix would not be present on the noun). One might therefore think that the phenomena we are dealing with receive a natural explanation by assuming a special semantics of the plural marker in 'Type Hungarian' that differs from the semantics of plurals in 'Type English'. I will, however, show that this cannot be the case.

For the sake of the argument, let us suppose Hungarian  $-(a)k/- (e)k/- (o)k/- (\ddot{o})k$  and its equivalents in the other languages have a meaning that differs from  $-s$  and its equivalents in 'Type English'. We would have to specify then what other semantics that might be. It is obvious that it is not the semantics of the distributive of the North-American languages. This is evidenced by the fact that the 'Type Hungarian' noun plural markers occur regardless of whether the combination of verb and argument denotes one or several events. Also note that they are obligatory where they occur; in particular, there are no plurality splits depending on the semantics of the noun. This indicates that number is grammaticalised as an inflectional category in 'Type Hungarian' languages, in contrast to a derivational status as described by Mithun.

If the semantics of 'Type Hungarian' plural differs from that of 'Type English', what other meaning would it then have? Could it be a functor over individuals that yields a collective interpretation? First, it is totally unclear just what the semantics would be precisely, and second, a semantics that deviates from standard plural analyses (such as Link 1983; Ojeda 1993) would only be justified if there were clear empirical evidence for it. For example, one would expect phenomena such as collectivity with the plural form, or transnu-

merality with the unmarked form (that is, the latter would in fact not be a singular; see 3.1 below), but this is not what we find in connection with the plural markers of 'Type Hungarian' languages.

Two interesting cases should be mentioned here that at first sight seem to contradict the claim just made, namely that we do not have to do with collectivity:

(i) The Hungarian plural suffix etymologically derives from a collective suffix. At first sight, this original function still seems to have a synchronic reflex, namely the suffix *-ék*, which attaches to proper names and denotes a group associated to the person referred to by the name. This construction indeed triggers plural agreement on the verb:

- (17) János-ék                sört    isz-nak.  
 János-collective    beer    drink-3pl  
 'János and his folks are drinking beer.'

However, unlike the plural suffix *-(a)k/-(e)k/-(o)k/-(ö)k*, which is subject to vowel harmony, the shape of the collective is phonologically invariant (note that in (17) it is suffixed to a stem with back vowels). Therefore, the collective *-ék* is a suffix of its own, contrasting to the plural suffix.<sup>8</sup> In other words, the semantics of the Hungarian plural suffix is not that of a collective (since this latter function is expressed by a different suffix), and therefore cannot explain the lack of plural agreement in the Hungarian noun phrase.

(ii) The Georgian plural suffix *-eb* used to function as a collective suffix in Old Georgian, according to Harris (1985). At that stage, there was a plural suffix also indicating case, for example, *-ta* for the dative. Modifiers exhibited overt number agreement with the noun, see (18e,f).

- (18) Old Georgian number agreement (Harris 1985, 194–96):

| singular: |                      | collective: |                                | plural: |                            |
|-----------|----------------------|-------------|--------------------------------|---------|----------------------------|
| (a)       | saxl-sa    mas       | (b)         | saxl-eb-sa    mas              | (c)     | saxl-ta    mat             |
|           | house-dat    df.sg   |             | house-coll-dat    df.sg        |         | house-dat.pl    df.pl      |
|           | 'the house'          |             | 'the houses'                   |         | 'the houses'               |
| (d)       | saxl-sa    did-sa    | (e)         | saxl-eb-sa    did-eb-sa        | (f)     | saxl-ta    did-ta          |
|           | house-dat    big-dat |             | house-coll-dat    big-coll-dat |         | house-dat.pl    big-dat.pl |
|           | 'big house'          |             | 'big houses'                   |         | 'big houses'               |

Note that the collective suffix *-eb* must formally be specified as singular, as is clear from the form of the agreeing determiner, see (18a,b) vs. (18c). At a later

<sup>8</sup> As an anonymous reviewer points out, the collective suffix *-ék* is the result of combining a 'pro-element' for a non-realised possessum, *-é*, with the plural suffix.

stage, then, *-eb* takes over the function of the original plural suffix *-ta*, still showing, however, overt agreement on all targets. In other words, agreement is still maintained after the reanalysis of *-eb* from collective to plural (now in terms of [+pl]), and is lost only later—obviously due to a change of the adjective from postnominal to prenominal. That is, the lack of plural agreement in Modern Georgian is not due to the ‘collective’ origin of the plural marker.

We can conclude, then, that the plural semantics of ‘Type Hungarian’ is the same as that of ‘Type English’. Consequently, in order to deal with the problems summarised at the end of section 1 we need a more principled account than one that simply stipulates a particular plural operator.

### 2.3. The questions to be answered

At this point it should be pointed out that ‘Type Hungarian’ plural constructions have occasionally been observed and commented on, most notably by Jespersen (1924, 208):

“In Magyar, [...] number is indicated in a secondary and not in a primary word, but only when a substantive is accompanied by a numeral. It is, then, put in the singular as if we were to say ‘three house’. This is termed ‘illogical’ by the eminent native linguist Simonyi: I should rather call it an instance of wise economy, as in this case any express indication of the plurality of the substantive would be superfluous.”

Thus, on the one hand, it is evident at some intuitive level that we have to do with language economy here, and therefore this construction type represents a perfectly natural rather than a typologically marked option. However, on the other hand, no theoretical analysis is known that would predict where exactly plural is explicitly marked, and, more intricately, how agreement should be handled formally in each case. The theoretical questions we are left with are therefore the following:

- How are the plural affixes formally excluded on modifiers, verbs, or on the noun itself, in the context of plurality?
- Is it possible to capture noun-phrase internal agreement by agreement in the formal sense of feature unification?
- What are the exact feature specifications of the categories involved?

The goal of the remainder of the paper is to look for an analysis of the facts that answers these questions. I will first sketch two different approaches that



have the advantage of utilising more conventional tools of linguistic description, and show why they fail to adequately account for the facts. Finally, I propose a constraint-based analysis under which plural marking is excluded by a constraint of morphological economy, thus building on Jespersen's intuition.

### 3. Why conventional solutions are bound to fail

The two alternative analyses that are worth being tried out due to their conceptual simplicity are one that is based on underspecification (3.1) and one that is based on selection (3.2).

#### 3.1. Unification and underspecification

The first analysis is based on the (now standard) conception of agreement as unification of features, as it is most explicitly performed in HPSG (Pollard-Sag 1994, chapter 2) and Wunderlich (1994). Moreover, the analysis makes use of the concept of underspecification, which is often used in linguistic theory for the representation of syncretism and unmarked forms. A further essential of the analysis is that 'Type Hungarian' differs from 'Type English' in that only in the latter type will all nouns or agreement targets that occur without their plural marker be specified as  $[-pl]$  per default for reasons of paradigmatic opposition.

The standard use of default value assignment for number inflection is briefly illustrated for 'Type English'. In this type, nouns that are projected into the syntax are fully specified for number. Therefore, the base form, which paradigmatically contrasts to a noun marked for  $[+pl]$ , has the specification  $[-pl]$ , where the latter results from a default at the lexicon-syntax interface (see, for example, Wunderlich-Fabri 1996). This gives rise to the following representations for English nouns and verbs (the predicate 'AGGR(egate)' used in the semantic formula of the representation in (19b) and throughout the paper is due to Pollard-Sag 1994 and stands for an aggregate individuation of referential objects, thus denoting the semantic function of plural):

- (19) (a) house:  $[+N, -V]^9$ ;  $\lambda x^{(-pl)} [HOUSE(x)]$   
 (b) houses:  $[+N, -V]$ ;  $\lambda x^{(+pl)} [HOUSE(x) \ \& \ AGGR(x)]$

<sup>9</sup> As an alternative to the traditional major category features  $[\pm N, \pm V]$ , Wunderlich (1996) provides a more principled approach to features classifying the four major categories. Since this matter is of little relevance for the present purpose I stick to the more common features throughout this paper.

- (20) (a) tumbles down:  $[-N,+V]$ ;  $\lambda x^{(-pl)}$  [TUMBLE-DOWN(x)]  
 (b) tumble down:  $[-N,+V]$ ;  $\lambda x^{(+pl)}$  [TUMBLE-DOWN(x)]

These specifications immediately account for the ungrammaticality of phrases such as \**these house* or \**two house*: the unification of  $[-pl]$  on the noun and  $[+pl]$  on the demonstrative or the numeral simply fails. The same holds for subject-verb agreement, hence \**The houses tumbles down*.

By contrast, these standard assumptions about default values are rejected for base forms in 'Type Hungarian' under the analysis sketched here. No default rule adding the specification  $[-pl]$  to the base form at the lexicon-syntax interface applies here. As a consequence, only nouns marked for plural have the same representation as 'Type English' nouns, whereas the base form of the noun (the 'singular' form) formally remains underspecified when inserted in the syntax:

- (21) Hungarian nouns:  
 (a) 'singular' noun: hajó:  $[+N,-V]$ ;  $\lambda x^{( )}$  [SHIP(x)]  
 (b) plural noun: hajó-k:  $[+N,-V]$ ;  $\lambda x^{(+pl)}$  [SHIP(x) & AGGR(x)]

As (21a) illustrates, the singular form of a noun is unspecified for the feature  $[pl]$ . In other words, a noun not explicitly marked for  $[+pl]$  will be underspecified with respect to number ( $[ ]$ ) under this analysis, and thus unification with a  $[+pl]$  form is possible. Similarly, an attributive adjective would be underspecified:

- (22) Hungarian attributive adjective: gyors:  $[+N,+V]$ ;  $\lambda x^{( )}$ [FAST(x)]

An analysis along these lines would correctly predict the grammatical examples of the data: the non-plural form of the adjective combines with a plural noun, as in *gyors hajók* (example (2) above), and a non-plural noun combines with numerals and quantifiers, as in *öt hajó* (example (4)).

However, there are some inevitable problems for such an analysis based on underspecification: while it accounts for the grammaticality of 'Type Hungarian' constructions, it is much too permissive, in that it fails to explain the ungrammaticality of combinations of  $[+pl]$  and  $[+pl]$ . Rather, it would allow for them since unification is just as possible here as in the case of  $[+pl]$  and  $[ ]$ . In other words, the analysis does not account for the impossibility of  $[+pl]$  adjectives with a plural noun as in (23a), or of  $[+pl]$  nouns with a numeral in (23b).

- (23) Hungarian: (a) \*gyors-**ak**    hajó-**k**    (b) \*öt    hajó-**k**  
                               fast-pl    ship-pl                    five    ship-pl

Similarly for DP-external agreement: if the non-plural form of a verb is unspecified, how come that in combination with a plural noun a [+pl] specification on the verb is grammatical (cf. the Georgian example (13a) above), but the unspecified ('singular') form is not (24a); and for a subject containing a numeral (cf. (13b)), [+pl] on the verb should be just as acceptable as the unspecified form, but this is excluded too, see (24b).

- (24) Georgian: (a) \*knut-eb-i    gorav-s    (b) \*sami knut-i    gorav-en  
                               kitten-pl-nom    roll**3sg**                    three kitten-nom    roll-**pl**

Finally, nothing excludes the combination of a [+pl] verb form with a [+pl] (instead of an unspecified) subject DP in Kurdish (see example (11) above).

In addition to these obvious empirical shortcomings, there is also a typological problem for an underspecification analysis: a noun form formally unspecified for number would suggest transnumerality; for example, an overt singular form in opposition to the unmarked form (which is found in many languages), or numeral classification (that is, counting is not available without overtly establishing a counting criterion). However, such phenomena are unknown in the languages in question, so there is no evidence whatsoever for transnumerality.

In sum, an underspecification analysis, which would be otherwise highly desirable because of its conceptual and representational simplicity, does not adequately capture the facts and must therefore be rejected.

### 3.2. Syntactic head selects number specification of its complement

Next to underspecification, there is a second strategy to account for the 'Type Hungarian' plural that comes to mind. It is based on the concept of a syntactic head selecting a particular specification of its complement. More specifically, the idea would be that a head specified as [+pl] selects a complement with the specification [-pl]. For the combination of noun and adjective, one would further assume that the attributive adjective is 'built in' into the composition as a complement of the noun. And for numeral constructions, the numeral would be the syntactic head of the noun phrase. Such an analysis is illustrated by the following categorial grammar style representations of Hungarian *hajó-k* 'ship-pl' and *öt* 'five'.

- (25) (a) hajó-k: [+N,-V,+MAX,+pl]/[+N,+V,+MAX,-pl];  
 $\lambda P \lambda x^{(+pl)}[(SHIP(x) \ \& \ AGGR(x) \ \& \ P(x))]$
- (b) öt: [+N,-V,+NUM,+MAX,+pl]/[+N,-V,+MAX,-pl];  
 $\lambda N \lambda x^{(+pl)}[N(x) \ \& \ CARD(x)=5]$

The noun in (25a) is, according to the assumption just mentioned, extended into a head selecting an adjective.<sup>10</sup> Formally, it is a functor that takes a [-pl] adjective phrase (to the right of the slash) and produces a noun phrase with the specification [+pl] (to the left of the slash). Similarly, the numeral in (25b) is represented as a head that takes a [-pl] noun and produces a maximal phrase with the specification [+pl].

This treatment would also extend to plural verbs in Kurdish, which were shown to combine with a singular subject. This is illustrated in (26) for the verb form *dixwîn-in* '(they) learn'.

- (26) dixwîn-in: [-N,+V,+MAX,+pl]/[+N,-V,+MAX,-pl];  $\lambda x^{(-1,-2,+pl)} [LEARN(x)]$

The entry states that a verb specified for plural selects a subject with the specification [-pl].

However, an account of the data along the lines sketched here is for several reasons no more successful than the underspecification analysis discussed in the preceding subsection. The first problem inherent to this account is obvious: it is highly technical and unintuitive. In particular, there is no plausible reason why a plural head should select a singular complement.

A second and more fatal problem is that although such an analysis might work for the simpler cases illustrated in (25) and (26), confronted with the more intricate facts it runs into a headedness dilemma for both noun phrases and clauses. By this I mean a conflict as to what is the head of the structure and what is the complement, insolvable on the grounds of serious linguistic assumptions. To start with, the numeral, though it seems to govern the number specification of the noun, cannot be the head of the noun phrase in most of the languages in question. This is evident from constructions such as the Archi example in (10) above, here repeated as (27):

<sup>10</sup> Treating noun modifiers as arguments of the head noun is not merely a construct of linguistic theory but is also justified by the morphology. Evidence comes from languages such as Persian and Kurdish, in which the extension of the noun into a head selecting a complement is overtly marked by the *ezafe*-suffix (see footnote 4). The issue is extensively dealt with in Ortmann (2000), where an account of the semantic and morphological aspects of the *ezafe* construction is provided.

(27) Archi (< Dagestan < NE Caucaus): (Kibrik 1994)

|                                                |         |          |              |           |           |         |
|------------------------------------------------|---------|----------|--------------|-----------|-----------|---------|
| yo-w                                           | zon     | L'annu-w | w-is         | q''le-w-u | dol:zu-w  | usdu    |
| dem-Isg                                        | pron1sg | love-Isg | Isg-pron.gen | two-Isg   | elder-Isg | brother |
| 'these two elder brothers of mine who love me' |         |          |              |           |           |         |

The example shows that the numeral *q''lewu* is placed between other modifiers, which casts serious doubt on its status as the head of the noun phrase. Instead, the construction is syntactically headed by a demonstrative, which is evidenced by its position and by syntactic and semantic assumptions about noun phrase structure (for more details see 4.3 below). Now one might argue that the numeral is the head of a local structure [*q''lewu [dol:zuw usdu]*], consisting of the numeral and the combination of adjective and noun, and consequently the former can select a [-pl] specification on the latter. However, even if this is accepted, it is hard to conceive how the numeral would formally select a specification of hierarchically higher constituents, namely the relative clause and the demonstrative. (Note that in the absence of the numeral, the other constituents would be marked for plural instead.) Hence, though the numeral determines the number value of its co-constituents, it is not the syntactic head of the structure.

Similarly, it is not the case that the verb form generally determines the number specification of the subject. Recall from section 1 that for all 'Type Hungarian' languages (except for Basque, on which see 4.5), plural verb forms do not combine with subject-NPs containing numerals. This is exemplified in (28) with verbal agreement in Georgian:

|                    |                            |          |     |                              |            |          |
|--------------------|----------------------------|----------|-----|------------------------------|------------|----------|
| (28) Georgian: (a) | knut-eb-i                  | gorav-en | (b) | sami                         | knut-i     | gorav-s  |
|                    | kitten-pl-nom              | roll-3pl |     | three                        | kitten-nom | roll-3sg |
|                    | 'The kittens are rolling.' |          |     | 'Three kittens are rolling.' |            |          |
|                    |                            |          |     | num                          | [-pl]      |          |
|                    | [+pl]                      | [+pl]    |     | [+pl]                        |            | [-pl]    |

A [-pl] verb does not necessarily combine with a [-pl] subject, but can also occur with a [+pl] 'numeral' subject (28b). On the other hand, a [+pl] subject without a numeral does not go with a [-pl] verb, but rather with [+pl], as in (28a). Thus the number value of the verb is determined by the structure of the subject-NP rather than vice versa. Hence the question arises if it is not the subject that would have to be treated as the head of the structure. However, the subject being the head of the clause is clearly at odds with all reasonable assumptions about phrase structure.

These problems can, of course, technically be handled in one stipulative way or other, but as long as the selectional analysis sticks to common-sense lin-

guistic assumptions about phrase structure, it runs into a headedness dilemma for both NPs and clauses.

In sum, the two analyses sketched in 3.1 and 3.2 show that an account of the ‘Type Hungarian’ plural that simply relies on underspecification or selection runs into problems when trying to explain the complexity of the facts. It is therefore obvious that a proper analysis of the ‘Type Hungarian’ plural is not a trivial task, and that a more flexible strategy is required.

#### 4. Economic number agreement

The analysis I propose in this section makes use of the idea of grammar consisting of violable constraints and, crucially, implements the notion of morphological economy. As I will show, it is thus capable of capturing the facts that were shown to be problematic for other conceivable solutions. The representations I propose will primarily be illustrated with data from Hungarian. A final subsection treats those aspects of other ‘Type Hungarian’ languages to which these representations do not immediately carry over.

The basic assumptions of the economic number agreement analysis are the following: (i) unification of agreement features, as in HPSG (Pollard–Sag 1994) or Wunderlich (1994) (in this regard the analysis does not differ from the two previous ones); (ii) violable constraints arranged in a language-specific ranking, as outlined in the framework of Optimality Theory (Prince–Smolensky 1993). Above all, I assume (iii) an economy principle, formally implemented as a constraint according to which number should not be realised redundantly within the determiner phrase (DP). This is stated in (29):

(29) PRINCIPLE OF ECONOMIC PLURAL MARKING (PEPL)

There is no more than one realisation of [+pl] within the DP.

In a framework that conceives of grammar as a set of constraints in a language-specific ranking, the requirement of unification of agreement features must also be implemented as a constraint:<sup>11</sup>

<sup>11</sup> As Jim Blevins (personal communication) suggests, the constraint PEPL can formally be implemented as a restriction on the mechanism of unification. The advantage would be that the occurrences of plural markers could be checked locally, rather than ‘being counted’ in a higher structure. However, this issue is not pursued any further here.

## (30) UNIFICATION:

Categories in a syntactic agreement relationship have compatible values with respect to agreement features (gender, number, person)

As far as the interaction with other constraints is concerned, I will assume UNIFICATION to be undominated with respect to any other relevant constraint. First, however, a further constraint that accounts for the mapping of semantic properties to morphological specifications is needed. In order to express this correspondence, Wunderlich (1994, 2f) assumes default rules stating that, for example, something that has the real-world property 'female' is usually treated as [+fem] in the morphology.<sup>12</sup> Similarly, a referential object individuated as an aggregate of single objects is usually treated as [+pl], which is expressed in (31).

$$(31) \text{ AGGREGATE}(x) \rightarrow x^{(+pl)} \quad \text{NON-AGGREGATE}(x) \rightarrow x^{(-pl)}$$

Such default rules transfer easily into an OT-constraint, which is given in (32):

## (32) MAP:

Semantic properties of individuals are reflected by a corresponding morphosyntactic feature specification in the output.<sup>13</sup>

I assume the ranking in (33) for 'Type Hungarian' languages, where the mapping constraint is dominated both by UNIFICATION and, crucially, by PEPL:

$$(33) \text{ Ranking for 'Type Hungarian': UNIFICATION, PEPL} \gg \text{MAP}$$

With these prerequisites we can now provide a detailed analysis of the various agreement constellations introduced in section 1.2.

#### 4.1. Nouns and adjectives in Hungarian

For the combination of nouns and adjectives in Hungarian, one further constraint is required that accounts for the difference between attributive and predicative adjectives. The traditional rule-based formulation would be a Feature Cooccurrence Restriction of the sort in GPSG and HPSG, possibly in

<sup>12</sup> See Pollard–Sag (1994) for a similar set of rules.

<sup>13</sup> In a more principled OT analysis, MAP could be conceived of as belonging to the family of the MAX-constraints as defined in Correspondence Theory (McCarthy–Prince 1995). However, in the absence of a precise theory of what exactly counts as the input for words and syntactic phrases, this issue must be left open.

terms of the rule in (34) which states that attributive adjectives are unspecified for the feature  $[\pm pl]$ , in other words bear no number specification.

(34)  $[+ATTR] \supset \neg[pl]$

To express this relation in the framework of OT, I assume a constraint that requires morphemes that fit in the morphosyntactic context to be realised on the surface. Formally, this is conceived of as a correspondence of semantic properties of individuals and inflectional affixes. If the language has a suffix that selects adjectives and specifies them for plural, it should be affixed in aggregate individuation contexts; if it is not affixed, this results in a violation of the constraint  $REALISE(\mu)$ . The constraint is stated in (35):

(35)  $REALISE(\mu)$ :

A semantic property of an individual is reflected by a corresponding inflectional affix of the morpheme inventory of the language ("morphemes are realised")

It follows from the definition that in languages where the adjective is generally uninflected,  $REALISE(\mu)$  is always satisfied trivially in adjective contexts, since in such a language there is no affix to realise. The crucial difference between  $REALISE(\mu)$  and MAP is that the latter relates semantic properties to feature specifications, whereas the former refers to the realisation of morphemes. As we will see, this distinction enables us to select between an uninflected form that is underspecified and one where a particular feature is instantiated. The relevant ranking for Hungarian is given in (36), with  $REALISE(\mu)$  being dominated by MAP:

(36) Complete relevant ranking for Hungarian:

$UNIFICATION, PEPL \gg MAP \gg REALISE(\mu)$

Given this inventory of constraints, we are able to answer the above question as to the feature specification of the categories involved. For our example, the intended reading is that of several fast ships, hence the input is the semantic representation of the conjunction of the predicates 'SHIP', 'FAST', and 'AGGR(egate)'; the output forms freely produced by the GEN function are surface forms associated with morphosyntactic specifications. The constraint ranking, then, gives us the evaluation of the relevant output candidates shown in (37).



(37) input:  $\lambda x$  [SHIP (x) & FAST(x) & AGGR(x)]

- |                                                                                                                                                                     |                                                                                                                                                                                                                                             |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>(a) gyors-ak    hajó-k<br/>             fast-pl    ship-pl<br/>             [+pl]    [+pl]<br/>                   [+pl]            * ! PEPL</p>                  | <p>(b) gyors-ak    hajó<br/>             fast-pl    ship<br/>             [+pl]    [-pl]<br/>                                   ✗            * ! UNIFICATION</p>                                                                            |
| <p>(c) gyors        hajó-k<br/>             fast        ship-pl<br/>                           [+pl]<br/>             [-pl]        ✗            * ! UNIFICATION</p> | <p>(d) gyors        hajó-k<br/>             fast        ship-pl<br/>             'fast ships'<br/>             [ ]        [+pl]<br/>                           [+pl]            ✗ * REALISE(<math>\mu</math>)</p>                           |
| <p>(e) gyors        hajó<br/>             fast        ship<br/>             [-pl]    [-pl]<br/>                   [-pl]            * ! MAP</p>                      | <p>(f) gyors        hajó<br/>             fast        ship<br/>             [ ]        [-pl]<br/>                           [-pl]            * ! MAP<br/>                                                   * REALISE(<math>\mu</math>)</p> |

The combination of [+pl] markers on both the adjective and the noun in (37a) violates the high-ranked principle PEPL, as there is more than one specification of [+pl] in the DP. The combination of a singular and a plural specification as in (37b) and (37c) gives rise to a violation of the equally high-ranked constraint UNIFICATION, indicated by '✗' in the gloss.

In (37d) the adjective is left unspecified for plural, hence this candidate faces a REALISE( $\mu$ ) violation. This violation, however, is not a fatal one, because (37d) respects the higher ranked constraints, in particular PEPL, since there is only one occurrence of [+pl]. Therefore, this candidate is the optimal one.<sup>14</sup>

Candidates (37e) and (37f) are also to be taken into account. For both of them the noun is specified as [-pl], hence they face no PEPL violation. However, though not being ungrammatical, these candidates can only be interpreted as singular, since they do not provide any cue for a 'plurality', or AGGREGATE, interpretation—no matter whether the adjective is specified

<sup>14</sup> An assumption being implicit so far is that a REALISE( $\mu$ )-violation is less severe for modifiers and specifiers than it is for heads, the intuition being that in the unmarked case, inflectional morphology is realised on the head of a phrase, but not necessarily elsewhere. Thus formally, a candidate \*gyors-ak [+pl] hajó [ ], with the plural suffix being left out on the noun rather than on the adjective, is excluded by a high-ranked constraint REALISE( $\mu$ , head) referring to the head of a local structure. However, for the sake of convenience, those constraints that are unviolated for all cases under consideration are omitted throughout (with the only exception of UNIFICATION). In particular, these are all ALIGN, MAX, DEP, and IDENT constraints.

for  $[-pl]$  as in (37e) or unspecified as in (37f). Formally, these candidates constitute fatal violations of the constraint MAP. As mentioned before, the idea is that if a linguistic sign (here: the noun phrase) contains certain conceptual properties of its denotate, the morphosyntactic feature specification should correspond to these properties. Since a plurality individuation requires the corresponding feature  $[+pl]$  in Hungarian, MAP, which dominates  $REALISE(\mu)$ , is violated. (Note that candidate (37f) in addition also violates  $REALISE(\mu)$ .)

Representations of the categories involved are given in (38) and (39). For nouns, the singular form is fully specified:

- (38) (a) singular: hajó:  $[+N, -V]$ ;  $\lambda x^{(-pl)}$  [SHIP (x)]  
 (b) plural: hajó-k:  $[+N, -V]$ ;  $\lambda x^{(+pl)}$  [SHIP (x) & AGGR(x)]

On the other hand, an attributive adjective, unlike its predicative counterpart (see (2e) above) is not inflected, hence is not specified for the feature  $[pl]$ . This is reflected in (39a,b) vs. (39c). Its combination with a plural noun, as the result of the composition of (38b) and (39c), is shown in (39d), with  $\langle +pl \rangle$  as the agreement index resulting from the unification of  $\langle \rangle$  and  $\langle +pl \rangle$ .

- (39) (a) predicative singular adjective (e.g., (A hajó) gyors.):  
 gyors:  $[+N, +V, -ATTR] \lambda y^{(-pl)}$  [FAST(y)]  
 (b) predicative plural adjective (e.g., (A hajó-k) gyors-ak.):  
 gyors-ak:  $[+N, +V, -ATTR] \lambda y^{(+pl)}$  [FAST(y) & AGGR(y)]  
 (c) attributive adjective (e.g., gyors (hajó-k)):  
 gyors:  $[+N, +V, +ATTR] \lambda y^{(\langle \rangle)}$  [FAST(y)]  
 (d) attributive adjective + noun (e.g., gyors hajó-k):  
 gyors hajó-k:  $[+N, -V, +MAX] \lambda z^{(+pl)}$  [SHIP(z) & AGGR(z) & FAST(z)]

Note that the fact that predicative adjectives are obligatorily inflected for plural also follows from the assumption of the  $REALISE(\mu)$  constraint: since the dominating constraint PEPL is vacuously satisfied in the DP-external context of subject and predicative, a candidate with plural on the adjective is preferred over a candidate without a plural suffix, the latter giving rise to a fatal  $REALISE(\mu)$  violation.

Finally, note that the representation of the attributive adjective (39c) closely resembles that of the underspecification analysis that was rejected in 3.1. The difference, however, is that under the present account we have an explanation of why  $[+pl]$  is excluded on the attributive adjective, namely in

terms of an economy constraint that outranks the realisation of affix material otherwise fitting in the morphosyntactic context.<sup>15</sup>

#### 4.2. Nouns and numerals

Let me now turn to the combination of nouns and numerals. Remember the ranking in (33) and (36), where MAP is dominated by the economy constraint PEPL and by UNIFICATION as well. Consequently, the evaluation of candidates for the combination of noun and numeral is as shown in (40) for Georgian:<sup>16</sup>

- (40) input:  $\lambda x$  [KITTEN(x) & AGGR(x) & CARD(x) = 3]
- |                                                                                                                                                      |                                                                                                                                                                     |
|------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>(a) sami knut-eb-i<br/>         three kitten-pl-nom<br/>         [+pl] [+pl]<br/>             [+pl]           * ! PEPL</p>                        | <p>(b) sami knut-i<br/>         three kitten-nom<br/>         [+pl] [-pl]<br/>                           ✗                   * ! UNIFICATION</p>                    |
| <p>(c) sami knut-eb-i<br/>         three kitten-pl-nom<br/> <br/>         [-pl] [+pl]<br/>                   ✗                   * ! UNIFICATION</p> | <p>(d) sami knut-i<br/>         three kitten-nom<br/>         'three kittens'<br/>         [-pl] [-pl]<br/>                   [-pl]                   ✗ * ! MAP</p> |

Again, plural on both items as in (40a) violates the economy constraint PEPL, whereas differing specifications as in (40b) and (40c) violate UNIFICATION. The winning candidate is (40d), where the numeral is not mapped to the morphological specification [+pl], in spite of its semantics indicating 'more than one', hence 'plurality'. Hence the specification [-pl] violates the MAP-constraint, but (40d) is still optimal according to the 'Type Hungarian' ranking, since

<sup>15</sup> Given the ban against plural on the attributive adjective, the question arises why the plural suffix is maintained at all on the predicative adjective, instead of a uniform behaviour of all adjectives. A functional explanation would be that Hungarian is a pro-drop language, thus predicates should be marked for agreement in order to license their arguments. Unlike other 'Type Hungarian' languages such as Georgian, Persian or Kurdish, Hungarian itself does not use a copula in the 3rd person present indicative; hence in predicative adjective constructions of this context, number agreement can be marked only on the adjective itself.

<sup>16</sup> Somewhat simplistically, I assume  $\Theta$ -identification to be the relevant compositional device for the semantic combination of noun and numeral here. 'CARD' represents a function that states the cardinality value of an individual.

it satisfies the higher-ranked constraints. (Note that unlike in (37e,f) an AGGREGATE interpretation in (40d) is available even though the specification is [-pl], namely because of the plurality semantics of the numeral.)

Consequently, the proper representation of 'Type Hungarian' numerals is as in (41), where the morphosyntactic number specification is [-pl] irrespective of the cardinality value:

- (41) Hungarian: öt: [+N,-V,+QUANT,-pl];  $\lambda x^{(-pl)}$  [AGGR(x) & CARD(x) = 5]  
 Georgian: sami: [+N,-V,+QUANT,-pl];  $\lambda x^{(-pl)}$  [AGGR(x) & CARD(x) = 3]

Numerals are thus correctly predicted to combine with singular nouns, as represented in (38a), yielding (42) as the representation for a 'Type Hungarian' noun phrase with a numeral:

- (42) öt hajó:  $\lambda z^{(-pl)}$  [(SHIP(z) & AGGR(z) & CARD(z) = 5)]

Note that the specification of 'Type Hungarian' numerals as [-pl] has an interesting theoretical implication. Under the present analysis, this specification is not simply stipulated in the lexicon but rather enforced by the grammar, more precisely by the ranking of constraints for the languages in question. According to recent work in Optimality Theory, this is a welcome result: analysing language-specific properties such as *do*-support as well as the lack of certain complementisers in English, and the behaviour and inventories of clitics in various Romance languages, Grimshaw (1997; 1998) arrives at the conclusion that the lexical inventory of functional categories in a language is governed by the constraints of the grammar, rather than vice versa. This is indeed what the present analysis yields for 'Type Hungarian' number agreement, where the existence of both [-pl] numerals and Hungarian uninflected attributive adjectives fall out as a consequence of the relative ranking of economy with respect to other constraints.

### 4.3. Plural agreement in Hungarian demonstratives: an apparent counterexample

In the syntax of the Hungarian noun phrase, there is exactly one environment where plural agreement is found, namely when a demonstrative pronoun is involved. The demonstrative has a singular and a plural form for both its proximate and its distal variant, viz. singular *ez* and plural *ezek*, and *az* and *azok*, respectively.

- (43) proximate: (a) ez a hajó (b) ezek a hajó-k  
 dem.sg df ship dem.pl df ship-pl  
 'this ship' 'these ships'
- distal: (c) az a hajó (d) azok a hajó-k  
 dem.sg df ship dem.pl df ship-pl  
 'that ship' 'those ships'

At first sight, this pattern seems to contradict the constraint ranking assumed so far, since it is obvious that two [+pl] specifications are involved in (43b,d). Apparently, at best an additional stipulation is needed. For example, one might want to revise the constraint PEPL as to state for a particular language if its domain is DP or only NP, the constituent selected by D. In fact, however, agreement with a demonstrative does not pose a problem for the economic plural analysis, no modification of PEPL nor any other stipulation is needed for Hungarian. To show this, a proper analysis of the syntax of the Hungarian noun phrase must be provided first.

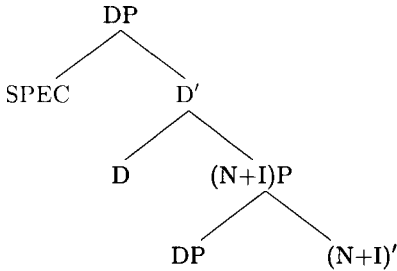
It is a crucial fact of the Hungarian noun phrase that, as in many other languages, the demonstrative is not in complementary distribution with the definite article (as is the case in 'Type English' languages), but rather the latter obligatorily follows the former. Compare (43) to the ungrammatical examples in (44):

- (44) (a) \*ez hajó (b) \*ezek hajó-k (c) \*az hajó (d) \*azok hajó-k  
 dem.sg ship dem.pl ship-pl dem.sg ship dem.pl ship-pl

The distribution of (43) and (44) shows that a structural representation of the Hungarian noun phrases has to accommodate the demonstrative next to the article. Nevertheless, its position is ignored in current proposals such as Szabolcsi (1994, 198). Szabolcsi suggests the representation in (45), which leaves no space for the demonstrative.<sup>17</sup>

<sup>17</sup> Accordingly, the demonstrative is not treated in any of Szabolcsi's examples that (45) is supposed to account for, although in a different context, its possible cooccurrences with quantifiers are explicitly mentioned.

(45) Structure of the Hungarian noun phrase according to Szabolcsi (1994, 198):



However, from the possible combination of the demonstrative with the definite article in (43) and also with the dative possessor in (46c) (the latter being an alternative to the center-embedded nominative possessor, compare (46a) to (46b)) we can conclude that a richer structure has to be assumed.

- (46) (a) az egyetemistá-k problémá-i  
 df student-pl problem-pl.por3  
 'the problems of the students'
- (b) az egyetemistá-k-nak a problémá-i  
 df student-pl-dat df problem-pl.por3  
 'the problems of the students'
- (c) az egyetemistá-k-nak ezek a problémá-i  
 df student-pl-dat dem def problem-pl.por3  
 'these problems of the students'

The cooccurrence of the dative possessor and the demonstrative clearly shows that the specifier position of D cannot accommodate all elements to the left of the determiner. I therefore assume that the Hungarian noun phrase is not simply a projection of D but rather has a projection above D, which is headed by the demonstrative.<sup>18</sup> I assume this projection to be the category Dem, which is the topmost functional category of the noun phrase, the hierarchy of syntactically relevant categories being Dem > D > AgrP(possess) > N. (The semantics of Dem is to fix the referent of the external argument of the noun relative to the utterance context.) However, in order to arrive at the proper syntactic composition, a further crucial fact has to be taken into account which shows that the relation between Dem and DP is not that of functional selection. As Kenesei–Vago–Fenyvesi (1998, 227f) observe, the repetition of a

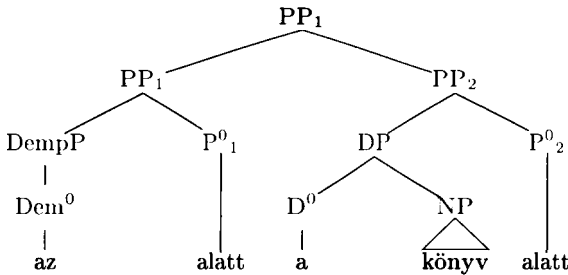
<sup>18</sup> Fabri (1993, 56–61) comes to the same result for Maltese.

morphological specification with demonstratives holds not only for plural, but also for the realisation of local relations:

- (47) (a) az alatt a könyv alatt (b) eb-ben a városban  
 dem under df book under dem-ine(ssive) df city-ine  
 'under that book' 'in this city'

The combination of a demonstrative and a postposition or a local case, respectively, renders a selection of DP by Dem<sup>0</sup>, or in other words, a structure [DemP [Dem<sup>0</sup> [DP]]], implausible: since postpositions take the entire noun phrase as their argument, their additional occurrence after the demonstrative in (47a) would be unexplained. Similarly, the case marker in (47b) would be expected to show up only once, namely on the lexical head; cf. example (16) above. The data can, however, be accounted for under an apposition structure which in English is rendered most closely as *under that one, under the book*. Thus, we are dealing with a PP the complement of which only consists of the demonstrative, where an appositional PP (here represented by 'PP<sub>2</sub>') with a DP as its complement is adjoined to PP<sub>1</sub>:

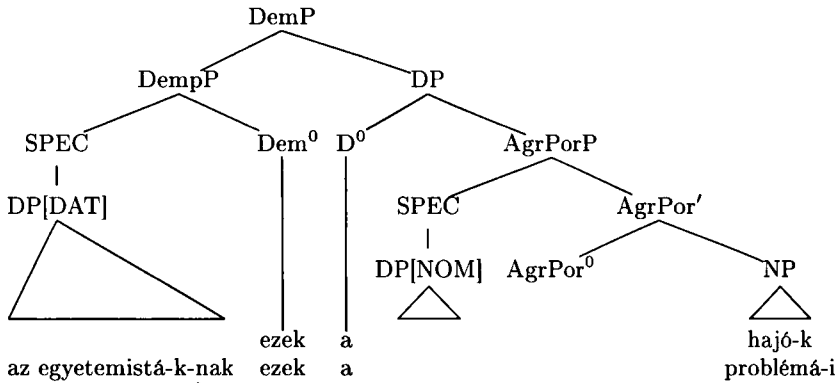
- (48) Hungarian PP with a DemP complement:



As far as the structure of a DemP not embedded in a PP is concerned, we can improve on Szabolcsi's analysis in terms of the apposition analysis and propose the structure in (49). The DP is adjoined to the topmost category DemP, which also hosts the dative possessor in constructions such as (46b). AgrPor corresponds to Szabolcsi's (N+I) and represents a projection for the nominative possessor argument, licensed by the agreement morphology on the noun (for more details irrelevant here see Szabolcsi's analysis).<sup>19</sup>

<sup>19</sup> I leave open the question of whether all Hungarian noun phrases, including those without a demonstrative, project to a DemP, or only those containing a demonstrative.

(49) Structure of the Hungarian noun phrase revised:



Given the essential part of the phrase structure in (49), the category  $\text{Dem}^0$  heading the noun phrases in (43), the agreement facts immediately follow from the definition of the constraint PEPL: as this constraint only refers to the level of DP, not to the higher domain DemP, PEPL cannot be violated in demonstrative–noun agreement. Quite obviously, at first sight this solution seems to be totally ad hoc: if the Hungarian noun phrase ultimately projects to a DemP, why should economy, and consequently the constraint PEPL, only refer to the DP, not to the entire DemP? However, the analysis seems far less stipulative if we take the apposition structure (49) more seriously. It is characteristic of appositions to redundantly repeat some grammatical information, while adding a more specific semantics. Given this, it is only natural that economy is not a relevant notion in such a construction. The obligatory use of a plural demonstrative with a plural noun is then predicted because otherwise there would be a failure of feature unification between head and adjunct. This is demonstrated by the following candidate evaluation:

- (50) (a) ez a hajó (b) ezek a hajó  
 dem det ship dem.pl det ship  
 [-pl] [] [-pl] [+pl] {} [-pl]  
 [-pl] \* ! MAP ✗ \* ! UNIFICATION
- (c) ez a hajók (d) ezek a hajók  
 dem det ship-pl dem.pl det ship-pl  
 [-pl] [] [+pl] [+pl] {} [+pl]  
 ✗ \* ! UNIFICATION [+pl] [+pl] ✗





Consequently, the proper representation for verbs must be as in (53): singular verbs are formally fully specified for number, namely as  $[-pl]$ , rather than being unspecified.

- (53) Georgian: (a) gorav-s:  $\lambda x^{(-1,-2,-pl)}$  [ROLL(x)]  
 (b) gorav-en:  $\lambda x^{(-1,-2,+pl)}$  [ROLL(x)]  
 Hungarian: (c) isz-ik:  $\lambda x^{(-1,-2,-pl)}$  [DRINK(x)]  
 (d) isz-nak:  $\lambda x^{(-1,-2,+pl)}$  [DRINK(x)]

With these representations, verbal agreement of a Hungarian subject noun phrase containing a demonstrative is immediately captured, given the account of noun phrases headed by a demonstrative as outlined in the previous subsection. Consider the data in (54).

- (54) (a) Ez a nagynéni sört isz-ik. (b) Ezek a nagynénik sört isz-nak.  
 dem df aunt beer drink-3sg dem.pl df aunt-pl beer drink-3pl  
 'This aunt drinks beer.' 'These aunts drink beer.'  
 (c) Ez az öt nagynéni sört isz-ik.  
 dem df five aunt beer drink-3sg  
 'These five aunts drink beer.'

The singular subject in (54a) combines with a singular verb, the plural subject in (54b), unlike the subject containing a numeral in (54c), combines with a plural verb. This is exactly what is expected under the present account, as any other combination would give rise to a fatal unification violation.

Finally, possessor agreement is to be considered, which also receives a straightforward explanation under this account. Consider the facts in (55):

- (55) (a) A nagynéni-k<sub>i</sub> ül-nek a  $\emptyset_i$  ház-uk<sub>i</sub>-ban.  
 df aunt-pl sit-3pl df pron house-por3pl-inessive  
 'The aunts are sitting in their (own) house(s).'
- (b) Az öt nagynéni<sub>i</sub> ül a  $\emptyset_i$  ház-á<sub>i</sub>-ban.  
 df five aunt sit.3dg df pron house-por3sg-inessive  
 'The five aunts are sitting in their (own) house(s).'
- (c) A nagynéni-k<sub>i</sub> ül-nek a  $\emptyset_i$  szék-ük<sub>i</sub>-ön.  
 df aunt-pl sit-3pl df pron chair-por3pl-locative  
 'The aunts are sitting on their (own) chairs.'
- (d) Az öt nagynéni<sub>i</sub> ül a  $\emptyset_i$  szék-é<sub>i</sub>-n.  
 df five aunt sit.3sg df pron chair-por3sg-locative  
 'The five aunts are sitting on their (own) chairs.'

A noun can be suffixed by a possessor marker, thus exhibiting agreement with an overt or non-overt pronoun denoting the possessor.<sup>20</sup> For better illustration, the non-overt pronoun is coreferential and hence coindexed, with the subject of the sentence in each of the examples in (55). If the pronoun cross-referenced by the possessor suffix refers to a noun phrase modified by a numeral, as in (55b,d), the possessor suffix on the noun is singular. (Note that as far as the number of the possessum is concerned, even in spite of the obvious reference to five chairs, in (55d) the singular form of the noun is used. The same form is preferred in (55c) where the plural form *szék-ei-k-en* is also acceptable though dispreferred, *-ei-* being the allomorph for possessed plural nouns. The use of the plural form in this example, however, gives also rise to the alternative reading ‘... each on their several chairs’.)

The distribution of pronominal possessor agreement in (55) is parallel to subject–verb agreement as described above; in fact, both verb agreement and possessor agreement are instances of noun-phrase-external agreement, in so far as the possessum is external with respect to the possessor phrase. Hence the representation of numerals as [–pl] and that of verbs in (53) extend to the representation of pronominal possessor agreement: only the form *ház-a* in (56a) (and similarly the form *szék-e*, as opposed to *szék-ük*), being specified as [–pl], can combine with a coindexed noun phrase containing a numeral.

- (56) (a) ház-a:  $\lambda y^{(+PRON, -1, -2, -pl)} \lambda x$  [HOUSE(x) & POSS(y,x)]  
 (b) ház-uk:  $\lambda y^{(+PRON, -1, -2, +pl)} \lambda x$  [HOUSE(x) & POSS(y,x)]

<sup>20</sup> It is important to note that a plural specification of the possessor marker as illustrated in (55a,c) is only found with pronominal, not with lexical possessors. This can be shown by the following contrastive pairs where (ib) is the pronominal variant of both (ia) and (ic), whereas (id) is that of e.g. *a nagynéni házában* ‘in the aunt’s house’:

- |                                                                                             |                                                                             |
|---------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------|
| (i) (a) a nagynénik ház-á-ban<br>df aunt-pl house-por3sg-inessive<br>‘in the aunts’ house’  | (b) az ő ház-uk-ban<br>df pron3sg house-por3pl-inessive<br>‘in their house’ |
| (c) az öt nagynéni ház-á-ban<br>df five aunt house-por3sg-inessive<br>‘in the aunts’ house’ | (d) az ő ház-á-ban<br>df pron3sg house-por3sg-inessive<br>‘in her house’    |

The plural of the possessor is not marked in combination with full NP possessors, as is clear from the invariant use of the singular form in (ia) and (ic). Rather, it is only marked in combination with pronominal possessors, as is clear from (ib) vs. (id) (and from (55a,c) vs. (55b,d) as well). Possessor agreement within the DP thus comprises only person, whereas number is only relevant with respect to an antecedent to be identified outside the DP, as in (55). I am grateful to Péter Siptár and an anonymous reviewer for pointing this out.

Summing up, the economic agreement analysis is capable of predicting the facts of ‘Type Hungarian’ plural constructions for all the agreement constellations involved: adjective–noun, numeral–noun, demonstrative–noun, and subject–verb. The essential tool of the analysis is an economy constraint ranked above the constraints MAP and REALISE( $\mu$ ).

#### 4.5. Residue: other plural constructions in ‘Type Hungarian’ languages

While most of the candidate evaluations and representations proposed so far hold for all ‘Type Hungarian’ languages, a few plural constructions are found for which a slight modification is necessary. Two of these constructions will be treated here briefly, in order to show that only minor language-specific properties are involved that can be implemented as additional constraints, rather than requiring a different ranking of the constraints relevant for our analysis.

To start with, Basque, as we have seen already, combines numerals with a singular noun and thus falls under the account given in 4.2 above. However, unlike other ‘Type Hungarian’ languages, it combines noun phrases containing a numeral with verbs marked for [+pl] rather than [–pl]. This is exemplified for subject agreement in (57a) and for indirect object agreement in (57b):<sup>21</sup>

- (57) (a) Pedro-ren lau etxe erre dira / ?da.  
 Pedro-gen four house burn.perf aux.3pl / aux.3sg  
 ‘Pedro’s four houses burnt down.’
- (b) Bost andre-ri eman diet / ?diot.  
 five woman-dat give.perf 3sg.abs-3pl.dat-1sg.erg / 3sg.abs-3sg.dat-1sg.erg  
 ‘I gave it to five women.’

This behaviour can be accommodated if we assume the following. With respect to the features under consideration here, the structure of a Basque noun

<sup>21</sup> The same pattern is found in subject–verb agreement in the Kabardian dialect of East Circassian, a North-West Caucasian language:

- (i) East Circassian (Colarusso 1989, 291):
- Pšaasa-əy-pʰə-r ma-a-d-a-ha(-r).  
 girl-num(eral connective)-four-abs 3-pres-sew-intr-pl(-pres)  
 ‘The four girls are sewing.’

Although the structure of the Circassian noun phrase differs in minor details, the analysis for Basque developed here straightforwardly applies to (i).

phrase is as in other 'Type Hungarian' languages, that is, a DP containing a numeral has the formal specification [-pl]. As for verbal agreement, Basque exhibits so-called semantic agreement, which is frequently found, for example, with collective nouns in many other languages. An example is (British) English, which shows systematic agreement mismatches in constructions such as *Her family are good singers*. Formally, this behaviour can be conceived of as specifying an additional index of agreement features according to the semantics (here: [+pl]), which is relevant for the syntactic composition, while the index of the morphologically overt specification is existentially bound.<sup>22</sup> The subject noun phrase of (57a), then, can be rendered as in (58) ('D' represents the definiteness operator).

(58)  $Dy^{(+pl)} \exists x^{(-pl)} [\text{HOUSE}(y) \ \& \ \text{AGGR}(x) \ \& \ \text{CARD}(y)=4 \ \& \ \text{POSS}(\text{Pedro},y) \ \& \ x=y]$

Under the present account, the introduction of a new index is necessitated by the requirement for semantically driven subject-verb agreement, which is a consequence of the constraint MAP dominating a further constraint, stating that in the unmarked case, unification of agreement features should be according to overt morphological specification.<sup>23</sup> The constraint and the ranking are proposed in (59) and (60), respectively.

(59) FORMAL AGREEMENT:

Overt morphological specifications of categories in a syntactic agreement relationship are compatible with respect to agreement features (gender, number, person)

(60) Ranking for Basque:

UNIFICATION, PEPL  $\gg$  MAP  $\gg$  FORMAL AGREEMENT

This ranking correctly predicts that (semantic) plural agreement on the verb is preferred over (formal) singular agreement in numeral constructions:

<sup>22</sup> See Pollard – Sag (1994, chapter 2) for an explicit account – in HPSG, and Ortman (1992) for a representation in the lambda calculus, as it is used throughout this paper.

<sup>23</sup> Note that this latter condition is not necessarily required by the constraint UNIFICATION; cf. the definition in (30) above. For the other 'Type Hungarian' languages FORMAL AGREEMENT is assumed to be ranked as high as not to interact with any of the other relevant constraints.



tal and idiolectal variation, often giving rise to optionality in the grammar, is expected in an area such as semantically motivated agreement mismatches.<sup>25</sup>

Another slight modification of the economic number agreement analysis developed so far is in order for Kurdish. As shown in (11) above, here repeated as (62), in Kurdish a subject noun phrase is not marked for plural if the verb is specified for plurality of the subject.

(62) Kurdish: (Akrawy 1982)

- |                                                       |                                                                       |
|-------------------------------------------------------|-----------------------------------------------------------------------|
| (a) Mirov hat.<br>man come.past.sg<br>'The man came.' | (b) Mirov(*-an) hat-in.<br>man(*-pl) come.past-3pl<br>'The men came.' |
|-------------------------------------------------------|-----------------------------------------------------------------------|

Obviously, in Kurdish the requirement for economic plural agreement is even stronger than in other languages, as one overt specification of plural suffices even for a domain outside of the noun phrase. Under the present account this means that the economy constraint relevant here refers to the whole clause, rather than only to the DP.

An interesting question is whether this constraint—which I call PEPL (Clause)—should be introduced as a constraint in addition to the constraint PEPL used so far, or if OT should allow for language-specific versions of the constraints. The former option seems more plausible under the commonly-held view of constraints being universal (and is in full accordance with the common practise in contemporary work on phonology).

In order to predict agreement with a suffixless noun in (62b) as grammatical, we must assume for Kurdish that the additional constraint is ranked at least as high as to dominate  $\text{REALISE}(\mu)$ , hence UNIFICATION, MAP, PEPL(Cause)  $\gg$   $\text{REALISE}(\mu)$ . This yields the following evaluation of candidates for verbal agreement:

<sup>25</sup> An alternative solution that might come to mind would be a ranking for Basque where UNIFICATION is dominated by MAP and would thus be violated by (55). However, recall from the discussion in 4.2 that the behaviour of numeral constructions requires the reverse ranking (UNIFICATION  $\gg$  MAP). As construction-specific re-ranking of constraints is highly suspicious on conceptual grounds, I reject this solution.

(63) input: COME (Dx [MAN(x) & AGGR(x)])

- |                                                                                                                                                                    |                                                                                                                          |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|
| <p>(a) Mirov-an hat-in.<br/>         mau-pl come.past-3pl<br/>         [+pl] [+pl]<br/>               [+pl] * ! PEPL(Clause)</p>                                   | <p>(b) Mirov-an hat.<br/>         man-pl come.past<br/>         [+pl] [-pl]<br/>                   ✗ * ! UNIFICATION</p> |
| <p>(c) Mirov hat-in.<br/>         man come.past-3pl<br/>         [-pl] [+pl]<br/>               ✗ * ! UNIFICATION</p>                                              | <p>(d) Mirov hat.<br/>         man come.past<br/>         [-pl] [-pl]<br/>               [-pl] * ! MAP</p>               |
| <p>(e) Mirov hat-in.<br/>         man come.past-3pl<br/>         'The men came.'<br/>         [ ] [+pl]<br/>               [+pl] ✗ * REALISE(<math>\mu</math>)</p> |                                                                                                                          |

A plural suffix on both the noun and the verb as in candidate (63a) violates PEPL(Clause). Differing specifications as in (63b) and (63c) give rise to a unification violation. In (63d) neither constituent is marked [+pl], hence the constraint MAP, which states the correspondence between semantics and morphosyntactic specification, is violated. Candidate (63e) violates REALISE( $\mu$ ), since the nominal plural suffix, though being semantically required, is not realised; instead, the form is unspecified for number. However, since REALISE( $\mu$ ) is dominated by all other constraints in question, (63e) is the winning candidate.<sup>26</sup>

What this means for the representation of Kurdish nouns is that, unlike nouns of other 'Type Hungarian' languages (cf. (38) above), the base 'singular' form is unspecified for number rather than being specified for [-pl]:

- (64) (a) 'singular' noun: mirov: [+N,-V];  $\lambda x^{( )}$  [MAN(x)]  
 (b) plural noun: mirov-an: [+N,-V];  $\lambda x^{(+pl)}$  [MAN(x) & AGGR(x)]

Given this representation, the ranking assumed so far can also be maintained for Kurdish.

<sup>26</sup> For the question of how a candidate \**Mirov-an* [+pl] *hat* [ ], with the verb rather than the noun being left unspecified, is excluded see footnote 14.



By introducing a variant of the economy constraint PEPL that refers to the clause, then, the economic agreement analysis can easily be extended to cover further cases such as Kurdish subject-verb agreement.<sup>27</sup>

## 5. Conclusion

I have shown in this paper that number agreement in 'Type Hungarian' languages, that is, languages with plural constructions without plural concord, can be analysed by feature unification. Thus, formally speaking, the lack of number agreement in 'Type Hungarian' is only apparent.

The essential tool of the analysis is the economy constraint PEPL, ranked above two other constraints: MAP, which requires a correspondence of semantic properties of individuals (here aggregate individuation) and agreement morphology ([+pl]); and REALISE( $\mu$ ), which requires affix material that fits into the morphosyntactic context to be realised. A crucial result of the analysis is that a non-default mapping of aggregate semantics to the morphosyntactic specification [-pl] is a typological option in order to respect formal agreement.

Conceptually simpler analyses, in particular an underspecification analysis, have been shown not to be able to capture the data. The results of this work should therefore be taken as further support for the concept of language-specific rankings of morphological constraints as part of the grammar. In particular, lexical properties such as numerals > 1 specified as [-pl], uninflected attributive adjectives in Hungarian, and Kurdish unspecified noun stems, can be explained as a consequence of the language-specific ranking of constraints.

<sup>27</sup> Additional evidence for the existence of PEPL(Clause) next to PEPL comes from Abkhaz (North-West Caucasian). In Abkhaz, verb-agreement with non-human subjects is parallel to Kurdish, that is, plural is realised only on the verb, not on the subject, as in (ia) (Hewitt 1989, 44ff):

- |                                 |                        |                 |            |
|---------------------------------|------------------------|-----------------|------------|
| (i) (a) A-gəgs <sup>v</sup> °əg | sər-bà-yt.             | (b) A-là(-k°à)  | bzəya-k°à. |
| df-beast                        | obj1sg-subj3pl-see-aor | df-dog(-pl)     | good-pl    |
| 'the beasts see me'             |                        | 'the good dogs' |            |
| (c) Pət-y°ə+k'                  |                        |                 |            |
| several                         | df-girl(-pl)           | tall-pl         |            |
| 'several tall girls'            |                        |                 |            |

Unlike Kurdish, however, Abkhaz (at least optionally, depending on the dialect) exhibits DP-internal plural agreement, that is, plural is also marked on modifiers such as adjectives (cf. (ib,c)) and relative clauses. In other words, economy only plays a role outside of the DP. The implication would be that this language ranks PEPL(Clause) above REALISE( $\mu$ ), which in turn dominates the DP-constraint PEPL in some dialects.

This result is in line with the view held most radically by Grimshaw (1998), namely that variation between languages in the lexical inventory is enforced by the grammar.

For those languages that in contrast to the 'Type Hungarian' languages do allow redundant plural marking in the DP, here referred to as the 'Type English', the constraints are ranked in a way that MAP and REALISE( $\mu$ ) dominate the economy constraint PEPL. However, what is satisfied throughout for both types, is the constraint UNIFICATION, requiring unification of agreement features. Hence the ranking of the relevant constraints for 'Type English' is the following:

(65) Ranking for 'Type English': UNIFICATION, MAP, REALISE( $\mu$ )  $\gg$  PEPL

It should be pointed out that the plural constructions of 'Type Hungarian' languages do not reflect a typologically marked option. Rather, 'Type English' languages are equally 'unnatural', in that morphological economy seems to be considered less relevant here, and its pattern of plural concord also calls for a thorough explanation in terms of what exactly outranks PEPL. The balance between these two choices of unmarkedness is reflected by the analysis in terms of a different ranking of constraints requiring morphological economy (PEPL) and explicitness (MAP, REALISE( $\mu$ )), respectively.

## References

- Akrawy, F.R. 1982. Standard Kurdish grammar. Published by the author.
- Amin-Madani, S. – Lutz, D. 1972. Persische Grammatik. Groos, Heidelberg.
- Aronson, H.I. 1991. Modern Georgian. In: Harris, A.C. (ed.): The indigenous languages of the Caucasus, Vol. 1: the South Caucasian languages, 219–312. Caravan, Delmar N.Y.
- Barnas, U.B. – Salzer, J. 1994. Lehrbuch der kurdischen Sprache. Wien.
- Carstens, V. 1991. The morphology and syntax of Determiner phrases in Kiswahili. Ph.D. Dissertation, UCLA, Los Angeles.
- Cerron-Palomino, R. 1976. Gramática quechua: Junin-Huanca. Ministerio de Educación/ Instituto de Estudios Peruanos, Lima.
- Colarusso, J. 1989. East Circassian (Kabardian dialect). In: Hewitt, G.B. (ed.): The indigenous languages of the Caucasus, Vol. 2: the North West Caucasian languages, 261–355. Caravan, Delmar N.Y.
- Fabri, R. 1993. Kongruenz und die Grammatik des Maltesischen (Linguistische Arbeiten 292). Niemeyer, Tübingen.

- Grimshaw, J. 1997. Projection, heads, and optimality. In: *Linguistic Inquiry* 28: 373–422.
- Grimshaw, J. 1998. Optimization and variation in the functional lexicon. Paper presented at “Lexicon in focus: International Conference of the SFB 282 ‘Theory of the Lexicon’”, Wuppertal, August 17–19, 1998.
- Harris, A.C. 1981. *Georgian syntax. A study in Relational Grammar*. (Cambridge studies in linguistics 33). Cambridge University Press, Cambridge.
- Harris, A.C. 1985. *Diachronic syntax: the Kartvelian case*. (Syntax and semantics 18). Academic Press, Orlando FL.
- Hewitt, G.B. 1989. Abkhaz. In: Hewitt, G.B. (ed.): *The indigenous languages of the Caucasus, Vol. 2: the North West Caucasian languages*, 37–88. Caravan, Delmar N.Y.
- Holisky, D.A. – Rusudan, G. 1994. Tsova-Tush (Bastbi). In: Smeets, R. (ed.): *The indigenous languages of the Caucasus, Vol. 4: North East Caucasian languages, Part 2. The three Nakh languages and six minor Lezgian languages*, 147–212. Caravan, Delmar N.Y.
- Jespersen, O. 1924/1968. *The philosophy of grammar*. Allen and Unwin, London.
- Kenesei, I. – Vago, R.M. – Fenyvesi, A. 1998. *Hungarian*. Routledge, London.
- Kibrik, A.E. 1994. Archi. In: Smeets, R. (ed.): *The indigenous languages of the Caucasus, Vol. 4: North East Caucasian languages, Part 2. The three Nakh languages and six minor Lezgian languages*, 297–367. Caravan, Delmar N.Y.
- Lewis, G.L. 1967. *Turkish grammar*. Clarendon, Oxford.
- Link, G. 1989. The logical analysis of plural and mass terms: a lattice-theoretical approach. In: Bäuerle R. – Schwarze, C. – von Stechow, A. (eds): *Meaning, use, and interpretation of language*, 302–23. de Gruyter, Berlin.
- McCarthy, J. – Prince, A. 1995. Faithfulness and reduplicative identity. In: Beckmann, J. – Urbanczyk, S. – Walsh, L. (eds): *Papers in Optimality theory*. University of Massachusetts occasional papers in linguistics 18, 249–384. GSLA, Amherst.
- Mithun, M. 1988. Lexical categories and the evolution of number marking. In: Hammon, M. – Noonan, M. (eds): *Theoretical morphology. Approaches to modern linguistics*, 211–34. Academic Press, San Diego.
- Ojeda, A. 1993. *Linguistic individuals*. CSLI Lecture Notes no. 3. Stanford, CSLI.
- Ortmann, A. 1992. Zur Auflösung von Merkmalskonflikten unter Kongruenz. *Arbeiten des Sonderforschungsbereichs 282 “Theorie des Lexikons”* No. 24. Heinrich-Heine Universität, Düsseldorf.
- Ortmann, A. 2000. *Morphologische Kategorien des Nomens: Merkmale, lexikalischer Strukturaufbau und Ökonomie*. Doctoral dissertation, Heinrich-Heine-Universität, Düsseldorf.
- Pollard, C. – Sag, I.A. 1994. *Head-driven Phrase Structure Grammar*. The University of Chicago Press, Chicago and London.
- Prince, A. – Smolensky, P. 1993. *Optimality Theory: constraint interaction in generative grammar*. Ms., Rutgers University and University of Colorado. (To appear, Cambridge MA, MIT Press).

- Szabolcsi, A. 1994. The Noun Phrase. In: Kiefer, F. – É. Kiss, K. (eds): *The syntactic sucture of Hungarian. (Syntax and semantics 27)*, 179–274. Academic Press, New York.
- Wunderlich, D. 1994. Towards a lexicon-based theory of agreement. In: *Theoretical Linguistics 20*: 1–35.
- Wunderlich, D. 1996. Lexical categories. In: *Theoretical Linguistics 22*: 1–48.
- Wunderlich, D. – Fabri, R. 1996. Minimalist Morphology: An approach to inflection. In: *Zeitschrift für Sprachwissenschaft 14*: 236–94.
- Zimmermann, I. 1998. Das dativische pronominale Klitikum in der DP-Struktur des Bulgarischen. In: Fabri, R. – Ortmann, A. – Parodi, T. (eds): *Models of inflection. (Linguistische Arbeiten 388)*, 119–33. Niemeyer, Tübingen.

Address of the author: Albert Ortmann  
Seminar für Allgemeine Sprachwissenschaft  
Heinrich-Heine-Universität  
Universitätsstr. 1  
D-40225 Düsseldorf  
Germany  
ortmann@phil.fak.uni-duesseldorf.de

## ON THE INTERACTION OF PARADIGMATIC AND SYNTAGMATIC STEM ALTERNATION IN ITALIAN CONJUGATION\*

VITO PIRRELLI AND MARCO BATTISTA

### Abstract

Evidence from Italian conjugation shows that the structure of a verb paradigm can be characterised formally in terms of a distribution of slots into a number of equivalence classes, or **set partition**, where each equivalence class is associated with a distinct Aronovian **basic stem type**. A few set partitions account for the structure of all Italian verb paradigms, whether regular or less regular. This well-behaved family of distributions tightly constrains stem allomorphy at an appropriate level of abstraction, independently of whether the origin of allomorphy is morpho-phonological or purely morphological, showing the superiority of the obtained generalisations over syntagmatic accounts. In the paper, we illustrate some formal properties of the discovered partitions, and emphasise interesting connections with morphological productivity, historical change and language learning.

### 1. Background

In languages with a rich inflection, different stems of the same verb are often formally related in a predictable, systematic fashion. In Latin conjugation, for instance, there exists a well-known non-fortuitous correspondence between the past participle stem and supine and future participle stems. The future active participle stem *amatur-* seems to derive from the past passive participle stem (*amat-*) through addition of *-ur-*. The relationship is merely formal, in that it appears not to be accompanied by any correspondence in meaning: there is no acceptable sense in which the meaning of the future active participle includes that of the past passive participle (Aronoff 1994). Moreover, the relationship proves to be independent of how regular the process of stem formation in question is. Even with instances of unpredictable stem functions, as in the

\* This paper is the outcome of a joint effort. However, for the specific concerns of the Italian Academy only, V. Pirrelli is responsible for sections 1, 2, 3, 8 and 9, and M. Battista for sections 4, 5, 6 and 7.

case of the irregular past participle *visus* 'seen', the formal correspondence with *visurus* still holds perfectly. An identical point can be made for supine stem forms, such as *amatu* or *visu*, and their meaning. The ancient grammatical tradition used to treat these cases through the statement of a parasitic relation holding over word forms as wholes, whereby the member of a paradigm (a grammatical word) seems to be formed directly on another member of the paradigm.

In an attempt to offer a modern adaptation of the ancient treatment, Matthews (1991) proposes to capture this relationship by means of a **metarule**, that is a generalisation over rules of stem formation. In the case at hand, the metarule should refer to the set of Latin stem formation rules for the past participle, and derive from them a corresponding set for the future participle and supine. Aronoff (1994) objects that there is little reason for taking the Latin past participle as the base stem form from which all others should be derived, since there is no evidence that any of these forms is semantically or morphologically primary with respect to any other. He suggests factoring out the common formal core shared by the forms in question (say *amat-*) and associate it with a purely morphological index, namely **third stem**, accounting for its distribution across the paradigm.<sup>1</sup> Aronoff claims that all stems enjoy this property of being without morphosyntactic meaning. For this reason they are said to be **morphomes**, that is purely morphophonologically defined constructs, closer to theme vowels than to inflectional endings. In Aronoff's view, not all stems are equal. Some stems are, in a sense, more interesting than others, depending on the extent to which they meet any of the following three criterial properties:

- (1) (a) stems are not meaningful
- (b) stems are the input of morphological realisation rules of a language and enjoy as such a special status, as independent parts of the morphological system
- (c) stems are formal functions whose output may vary considerably according to the verb to which they apply

Aronoff claims that only the three traditionally recognised Latin **basic stem types**, namely the **present stem**, the **perfect stem**, and what he calls the **third**

<sup>1</sup> The idea that the past participle stem can be used as an independent basis for word formation processes other than purely inflectional ones is first entertained in Scalise (1981) in connection with Italian deverbal derivatives of the type *X+ione* (such as *formazione*, *descrizione*, *dispersione*, etc.) where *X* is the past participle stem of the verb base.

stem (used, as we saw, as a basis for the formation of past participle, supine and future participle forms) satisfy (1a-c) above to the full.

Pirrelli and Battista (1998) delve into some interesting formal properties falling out of Aronoff's basic stem types. They observe that the establishment of an inventory of basic stem types in the conjugation system of a language is equivalent to inducing a one-to-one mapping between variables ranging over stem alternants on the one hand, and classes of paradigmatic slots on the other hand, as illustrated in the following table, for Latin verbs in the active voice:

Table 1

| Latin Indexing Schema (active voice) |                                                                                                                          |
|--------------------------------------|--------------------------------------------------------------------------------------------------------------------------|
| $S_1$                                | $\leftrightarrow A = \{\text{prs\_i, impf\_i, fut\_i, prs\_s, imp\_s, prs\_imp, fut\_imp, prs\_g, prs\_prt, prs\_inf}\}$ |
| $S_2$                                | $\leftrightarrow B = \{\text{prf\_i, fprf\_i, pprf\_i, prf\_s, pprf\_s, prf\_inf}\}$                                     |
| $S_3$                                | $\leftrightarrow C = \{\text{pft\_prt, fut\_prt, fut\_inf, sup}\}$                                                       |

The schema enforces what Stump (1995) calls **stem indexing**, namely the assignment of an index to a stem, for the latter to be appropriately selected by a morphological realization rule. For example, it says that the stem  $S_1$  is required to be found as a basis for the formation of all present indicative forms (prs\_i), all imperfect indicative forms (impf\_i), etc. By the same token, the stem  $S_2$  is common to all forms of the perfect indicative (prf\_i), of the future perfect (fprf\_i), of the pluperfect (pprf\_i), etc.

Let  $P$  be the set of all paradigmatic slots making up the active paradigm of a non-defective Latin paradigm. It can be shown that, under normal circumstances, for  $A$ ,  $B$  and  $C$  defined as in Table 1,  $A \cap B = \emptyset$ ,  $A \cap C = \emptyset$  and  $B \cap C = \emptyset$ , where ' $\emptyset$ ' indicates the empty set. This, together with the further statement that the union of  $A$ ,  $B$  and  $C$  yields back the set  $P$  can be restated more formally by saying that the set  $\{A, B, C\}$  is a **partition** of  $P$ . We can eventually describe an indexing schema as a function from a set of basic stem types onto sets of paradigmatic slots as follows:

$$(2) \quad F: \{S_i\} \rightarrow 2^{|P|}$$

Crucially,  $F(S_i)$  induces a partition of  $P$ . Accordingly, Aronoff's basic stems can be seen as pointers to **partition classes** (PCs) of paradigmatic slots. This way of looking at them throws in sharp relief the systematic distribution of basic stem types across the paradigm, independently of considerations about their form.

If each single slot in the Latin verb paradigm is assigned an appropriate indexed stem variable according to the indexing schema of Table 1, one obtains a **distribution schema** such as the one in Table 2 overleaf (limited again to active forms only) where examples are given for two Latin verbs, a regular one, AMO 'love' (example1), and a less regular one, SONO 'play' (example2).

*Table 2*  
A morphomic distribution schema of Latin conjugation  
Finite forms

| TENSE/MOOD                | BASIC STEM     | EXAMPLE1  | EXAMPLE2  |
|---------------------------|----------------|-----------|-----------|
| present indicative        | S <sub>1</sub> | amo       | sono      |
| imperfect indicative      | S <sub>1</sub> | amabam    | sonabam   |
| future indicative         | S <sub>1</sub> | amabo     | sonabo    |
| present subjunctive       | S <sub>1</sub> | amem      | sonem     |
| imperfect subjunctive     | S <sub>1</sub> | amarem    | sonarem   |
| present imperative        | S <sub>1</sub> | ama       | sona      |
| future imperative         | S <sub>1</sub> | amato     | sonato    |
| perfect indicative        | S <sub>2</sub> | amavi     | sonui     |
| future perfect indicative | S <sub>2</sub> | amavero   | sonuero   |
| pluperfect indicative     | S <sub>2</sub> | amaveram  | sonueram  |
| perfect subjunctive       | S <sub>2</sub> | amaverim  | sonuerim  |
| pluperfect subjunctive    | S <sub>2</sub> | amavissem | sonuissem |

Non-finite forms

| TENSE/MOOD         | BASIC STEM     | EXAMPLE1      | EXAMPLE2       |
|--------------------|----------------|---------------|----------------|
| gerund/gerundive   | S <sub>1</sub> | amandum       | sonandum       |
| present participle | S <sub>1</sub> | amans         | sonans         |
| perfect participle | S <sub>3</sub> | amatus        | sonitus        |
| future participle  | S <sub>3</sub> | amaturus      | soniturus      |
| future infinitive  | S <sub>3</sub> | amaturus esse | soniturus esse |
| perfect infinitive | S <sub>2</sub> | amavisse      | sonuisse       |
| present participle | S <sub>1</sub> | amans         | sonans         |
| supine             | S <sub>3</sub> | amatum        | sonitum        |
| present infinitive | S <sub>1</sub> | amare         | sonare         |

It is important to note that identity of an index across two or more slots of Table 2 is not to be understood as implying surface formal identity of the



corresponding stems. For example, the imperfect indicative stem is formally distinct from the present indicative stem, in spite of their both being associated with  $S_1$  in Table 2.<sup>2</sup> Still both these stems are construed on the basic stem type  $S_1$ , and this captures a nonnegligible amount of intraparadigmatic formal redundancy. Similarly, the pluperfect stem is built upon  $S_2$ , the supine stem upon  $S_3$ , etc. Barring possible phonological readjustments, this morphomic redundancy holds consistently in the entire Latin paradigm.

## 2. Paradigmatic and syntagmatic stem alternation

The qualification “barring possible phonological readjustments” is important in this context: underlying phonemic identity of two stem forms is often blurred by phonological changes triggered by an embedding phonotactic context (e.g., the ensuing suffix). Let us consider an Italian example, the present indicative active of RISCHIARE ‘risk’ in Table 3.

Table 3

| PRESENT INDICATIVE |          |           |             |             |             |
|--------------------|----------|-----------|-------------|-------------|-------------|
| SINGULAR           |          |           | PLURAL      |             |             |
| 1                  | 2        | 3         | 1           | 2           | 3           |
| /’riskjo/          | /’riski/ | /’riskja/ | /ris’kjamɔ/ | /ris’kjate/ | /’riskjano/ |
| $S_1$              | $S_2$    | $S_1$     | $S_2$       | $S_1$       | $S_1$       |

The paradigm here shows two surface stem forms, /riskj/ and /risk/, considerably similar phonologically. This similarity has a natural phonological explanation: glide assimilation, rule (3). The distribution of the two alternants across the paradigm can be characterised as follows: /riskj(a)/ ( $S_1$ )<sup>3</sup> accompanies all endings except those beginning with /i, j/ (i.e. second person singular,

<sup>2</sup> Generally speaking, a stem form associated with any  $S_i$  can possibly undergo a phonological change as a result of either of the following facts: (i)  $S_i$  is input to a further stem formation rule, or (ii)  $S_i$  is trailed after by a termination triggering a particular phonological readjustment.

<sup>3</sup> Parentheses around the thematic vowel (/riskj(a)/) indicate that the vowel of  $S_1$  is not always overtly realised. This means that, for the purposes of base stem identification, both radical and thematic base forms are taken to be the same thing. This makes sense linguistically, as deletion of a thematic vowel is always predictable from the context of an ensuing suffix (Scalise 1983; Matthews 1991).

and first person plural). Phonologically, all surface stem forms and their distribution can be explained if one assumes that the following phonological rule (due to Scalise 1983) is operative in Italian:

(3) GLIDE ASSIMILATION

$$[j] \rightarrow \emptyset \left/ \begin{array}{c} [i] \\ [j] \end{array} \right.$$

This evidence clearly disfavours the hypothesis that the distribution of /riskj/ and /risk/ be captured paradigmatically, i.e. by means of a morphomic indexing schema. It would be reasonable to posit the existence of one underlying stem form only, namely  $S = /riskj(a)/$ , whose surface variant /risk/ is accounted for as the result of the application of glide assimilation in the appropriate context. This range of stem alternants is traditionally referred to as syntagmatic (Zwicky 1990).

Pirrelli and Battista (1998) describe in some detail the nature and distribution of stem alternation in Italian conjugation. They observe that, if one excludes cases of stem alternants which are mutually related through the statement of fully productive phonological rules of Italian such as glide assimilation, the paradigmatic distribution of all other cases is captured by the following Overall Distribution Schema (ODS).

*Table 4*  
Overall Distribution Schema (ODS)

| FINITE FORMS         |                |                |                |                |                |                |
|----------------------|----------------|----------------|----------------|----------------|----------------|----------------|
|                      | SINGULAR       |                |                | PLURAL         |                |                |
|                      | 1              | 2              | 3              | 1              | 2              | 3              |
| Present Subjunctive  | S <sub>2</sub> | S <sub>2</sub> | S <sub>2</sub> | S <sub>4</sub> | S <sub>4</sub> | S <sub>2</sub> |
| Present Indicative   | S <sub>2</sub> | S <sub>3</sub> | S <sub>3</sub> | S <sub>4</sub> | S <sub>1</sub> | S <sub>2</sub> |
| Imperfect Indicative | S <sub>1</sub> | S <sub>1</sub> | S <sub>1</sub> | S <sub>1</sub> | S <sub>1</sub> | S <sub>1</sub> |
| Imperfect Subj.      | S <sub>1</sub> | S <sub>1</sub> | S <sub>1</sub> | S <sub>1</sub> | S <sub>1</sub> | S <sub>1</sub> |
| Past Absolute        | S <sub>5</sub> | S <sub>1</sub> | S <sub>5</sub> | S <sub>1</sub> | S <sub>1</sub> | S <sub>5</sub> |
| Present Imperative   | —              | S <sub>3</sub> | S <sub>2</sub> | S <sub>4</sub> | S <sub>1</sub> | S <sub>2</sub> |
| Present Conditional  | S <sub>6</sub> | S <sub>6</sub> | S <sub>6</sub> | S <sub>6</sub> | S <sub>6</sub> | S <sub>6</sub> |
| Future Indicative    | S <sub>6</sub> | S <sub>6</sub> | S <sub>6</sub> | S <sub>6</sub> | S <sub>6</sub> | S <sub>6</sub> |

| NON-FINITE FORMS   |                |
|--------------------|----------------|
| Present Gerund     | S <sub>1</sub> |
| Present Participle | S <sub>1</sub> |
| Past Participle    | S <sub>7</sub> |
| Present Infinitive | S <sub>8</sub> |

The schema is obtained by selecting, for each array of tense-mood-voice properties (e.g., present indicative active, imperfect subjunctive active etc.), the attested distribution schema with the maximum number of stem alternants (stem indices). The following remarks are particularly relevant here:

- (1) ODS is abstract: there is no attested Italian verb exhibiting a different stem form in each of the eight partition classes. The maximum number of stem alternants witnessed in Italian conjugation is found in the paradigm of the verb *DOLERE* 'hurt', which shows six different instances of such variation<sup>4</sup> (with stem alternants S<sub>7</sub> and S<sub>8</sub> being replaced by S<sub>1</sub>).
- (2) ODS covers: (a) all cases of paradigmatic stem allomorphy of the sort considered by Aronoff in his analysis of Latin conjugation; (b) all stem alternants which have traditionally been analysed in terms of unproductive phonological processes, or *minor rules* (Lightner 1968; Hudson 1974).
- (3) all distributional schemata of Italian verbs, both regular and irregular ones, (with the only (partial) exception of eight highly suppletive verbs<sup>5</sup>) **are derived from ODS through set union of existing partition classes**, in accordance with the 'Schema Transition Hypothesis' (Pirrelli-Battista 1998); this is tantamount to saying that ODS can be turned into any attested schema by simply **re-indexing** one or more partition classes.

Point (2) is intriguing for its suggestion that paradigmatic and syntagmatic constraints over stem alternation interact in some non-trivial way. This should not be thought of as somehow undesirable; it is normal to find an interplay of separate factors which may in particular cases conflict with each other, in others reinforce each other, and in still others be independent of each other. The investigation of what determines the trading relation between these two sometimes conflicting dimensions is a difficult task. In the early 80's, Kiparsky (1982)

<sup>4</sup> The basic stems in question are: /dol/ (S<sub>1</sub>), /dɔlg/ (S<sub>2</sub>), /dwɔl/ (S<sub>3</sub>), /doɫ:/ (S<sub>4</sub>), /dɔls/ (S<sub>5</sub>), /dɔr:/ (S<sub>6</sub>).

<sup>5</sup> The eight truly exceptional verbs are: *AVERE* 'have', *ESSERE* 'be', *ANDARE* 'go', *DARE* 'give', *FARE* 'do, make', *STARE* 'stay, be', *DIRE* 'say, tell', *SAPERE* 'know'.

suggested that this investigation was nonetheless necessary if the notion of paradigmatic conditions was to yield precise predictions about the direction of possible historical language change. He hastened to add, however, that “it requires investigation into the structure of paradigms at a level which generative grammar at present may not be ready for”. More recently, Kiparsky (1996) points out that the morphological distribution of phonological processes, as well as the phonological distribution of allomorphs, are to be expected somehow: there is no reason to question the status of phonological and morphological stem alternation on this basis. Still the situation depicted in (2) should give us pause. We will return to it in sections 4, 7 and 8 of this paper.

Point (3) has interesting implications for language acquisition. The eight partition classes identified through ODS above define the maximum range of stem variability an Italian speaker is exposed to. This means that a speaker is in a position to reconstruct the entire Italian paradigm of any verb correctly, after hearing eight forms of this paradigm only, provided that each of these forms belongs to a different partition class of ODS. Care should be taken not to interpret this as implying either of the following: (i) that a learner of Italian inflectional morphology actually engages in looking for forms belonging to the eight partition classes identified here; (ii) that (s)he would eventually settle on establishing exactly eight partition classes. Assumption (i) is obviously unrealistic: it is highly unlikely that a learner acquainted with a verb *V* makes no generalisation about its inflectional behaviour until (s)he is exposed to eight different forms of *V*, each belonging to different partition classes. There is abundant psycholinguistic evidence that **imperfect learning** plays a prominent role in both directing human verbal behaviour and justifying language change. As to (ii), it is reasonable to assume that the actual number of partition classes which are eventually established by a mature human learner may vary depending on (a) random differences in the order and frequency of the evidence a learner is exposed to; (b) trade-off solutions between settling on the optimal classification strategy (eight classes make the exact prediction for any verb, whether regular or irregular) and choosing a suboptimal and less costly strategy which requires fewer parameters (e.g., fewer partition classes), but it nonetheless covers the vast majority of verb types in Italian.<sup>6</sup> In turn, the range of such trade-off solutions may change considerably depending on factors

<sup>6</sup> The so-called **Minimum Description Length** (MDL) approach provides a formal basis for evaluating the costs involved in the process of hypothesis selection. In particular, the MDL principle says: choose the hypothesis  $h_{MDL}$  minimising the following quantity:

$$-\log_2 P(D|h) - \log_2 P(h)$$

such as the conservative influence of the established norm, and the different stages in the learner's developing grammar.

In the remainder of this paper, we intend to suggest that the weaponry of formal descriptive notions introduced above sheds considerable light on the nature of this interaction. The fine-grained articulation of a paradigm into partition classes does not only have the merit of establishing a purely morphomic basis for the study of stem alternation but it also provides scholars with a fruitful, independent dimension of grammatical analysis. As will become clearer in the following pages, if one ignores the descriptive potential of this further dimension, one is left with the uncomfortable alternative of deciding, given the formal redundancy shared by two stem alternants, whether this should be accounted for in terms of (major/minor) productive phonological processes, or rather listed in the lexicon. Investigation of the formal properties of partition classes opens up another intriguing alternative account: **paradigmatically-governed productivity**.

### 3. The nature of ODS in Italian

Partition classes in Italian ODS (Table 4) can be seen as the result of the interaction of a variety of factors. First, some classes reflect, to an extent, the nature of the embedding phonotactic context. In particular: (1) the quality of the suffix vowel immediately following the stem (front vs. back), (2) the position of stress relative to the stem (stressed vs. unstressed). In other cases, partition classes are purely morphomic. In yet some other cases, they reflect the interaction of both morphomic and phonological factors. Here follows a classification of each partition class according to their defining properties:

---

Note that  $P(D|h)$  expresses the probability of  $h$  predicting the right classification of each verb type  $V \in D$ , where  $D$  is the training evidence to which the learner is exposed. Note further that  $\log_2 P(D|h) = 0$  when  $P(D|h) = 1$ , that is when the hypothesis  $h$  predicts the right class with certainty in all cases. On the other hand  $\log_2 P(h)$  is minimised when the parameters required by the hypothesis  $h$  have a high probability to be met during training. Certainly, being exposed to eight verb forms belonging to eight different partition classes is less likely (and thus more costly) than—say—being exposed to just two very frequent verb forms. MDL provides a way of trading off hypothesis complexity for the number of errors committed by the hypothesis. Note finally that, for what we said so far, there is no verb for which each  $S_i$  will be associated with a different stem. For most irregular verbs, only two or three different stems are found. With regular verbs, the same stem  $S_1$  is found in each partition class.

- (4) S<sub>1</sub>: default stem alternant  
 S<sub>2</sub>: stressed stem AND ensuing back vowel  
 S<sub>3</sub>: stressed stem AND ensuing front vowel  
 S<sub>4</sub>: unstressed stem AND ensuing glide  
 S<sub>5</sub>: stressed stem AND past absolute  
 S<sub>6</sub>: future indicative OR present conditional  
 S<sub>7</sub>: past participle  
 S<sub>8</sub>: infinitive

Three types of evidence can be offered in support to the analysis of S<sub>1</sub> as the default stem alternant. First, S<sub>1</sub> is by far the most widespread of all stem indices in ODS. Secondly, S<sub>1</sub> is always taken as input of regular stem functions. Thirdly, S<sub>1</sub> is always trailed after by regular (or weak) inflectional terminations. Finally, in the vast majority of cases, reduction in number of stem indices results in re-indexing a non-default index S<sub>i</sub> as S<sub>1</sub> (as a corollary of the Schema Transition Hypothesis).

It should be observed that S<sub>2</sub>, S<sub>3</sub> and S<sub>4</sub> are defined in purely phonological terms. All remaining partition classes are either morphomic (S<sub>6</sub>, S<sub>7</sub>, S<sub>8</sub>), or a combination of morphomic indexing and phonological factors (S<sub>5</sub>). In the following sections we give two reasons for judging S<sub>2</sub>, S<sub>3</sub>, S<sub>4</sub> and S<sub>5</sub> morphomic in spite of their correlation with clear phonological patterns.

### 3.1. Phonologically-governed distribution of suppletive stems

The main argument here is the existence of phonologically conditioned suppletives; they comply with the distributional classes indexed as S<sub>2</sub>, S<sub>3</sub>, S<sub>4</sub> and S<sub>5</sub>, but there exists no motivated phonological rule which can justify their surface form given the phonological context. Consider, for example, the past absolute indicative of MUOVERE 'move':

Table 5

| PAST INDICATIVE |                |                |                |                |                |
|-----------------|----------------|----------------|----------------|----------------|----------------|
| SINGULAR        |                |                | PLURAL         |                |                |
| 1               | 2              | 3              | 1              | 2              | 3              |
| /ˈmɔːsi/        | /mwoˈvesti/    | /ˈmɔːɛ/        | /mwoˈvem:o/    | /mwoˈveste/    | /ˈmɔːɛro/      |
| S <sub>5</sub>  | S <sub>1</sub> | S <sub>5</sub> | S <sub>1</sub> | S <sub>1</sub> | S <sub>5</sub> |

Here the distribution of two nearly suppletive stems, namely /mwov/ and /mɔːs:/, hardly correlates with any systematic variation of morphosyntactic

feature content. Observe further, however, that the stem /mwov/ (S<sub>1</sub>) is selected in the past absolute paradigm only when stress falls on the inflectional ending. On the other hand, an unstressed inflectional termination always follows the stem /mɔs:/ (S<sub>5</sub>). According to most scholars, this is a good example of what Carstairs (1990) dubs ‘phonologically-conditioned suppletion’, a phenomenon which Stump (1995) shows being common to languages other than Italian. In our terminology, phonologically-conditioned suppletions are an example of how unpredictability in the output of a stem function can correlate with a phonologically-predictable distribution.

### 3.2. Morphomic distribution of allegedly phonological processes

This sort of evidence is partitioned into two distinct classes: (i) underapplication and (ii) overapplication of alleged phonological processes. Underapplication takes place in cases of overregularization of present subjunctive forms of second conjugation verbs, whereby these verbs are wrongly inflected through first conjugation present subjunctive endings (see Table 6, where a question mark precedes cases of overregularization).

Table 6

| PRESENT SUBJUNCTIVE |         |         |            |            |           |
|---------------------|---------|---------|------------|------------|-----------|
| SINGULAR            |         |         | PLURAL     |            |           |
| 1                   | 2       | 3       | 1          | 2          | 3         |
| /vada/              | /vada/  | /vada/  | /an'djamo/ | /an'djate/ | /vadano/  |
| ↓                   | ↓       | ↓       |            |            | ↓         |
| ?/vadi/             | ?/vadi/ | ?/vadi/ |            |            | ?/vadino/ |

Such a shift of conjugation class has the immediate effect of fronting verb stems with the triggering environment /i/. The resulting outcome, however, always contradicts phonological predictions: in all these cases the stem preserves an S<sub>2</sub> stem alternant (as attested in the first person singular of present indicative): so we have ?/kreskino/, ?/di'rigino/ instead of the expected \*/kref:ino/, \*/di'riđino/.

The case of overapplication of minor rules is exemplified by the past participle forms of four subregular second conjugation *isc-* verbs such as CONOSCERE, CRESCERE, MESCERE and PASCERE: /kono'f:uto/, /kre'futo/, etc. Overt phonological evidence would predict /konos'kuto/, /kres'kuto/, etc. Appeal

to an underlying \*/konos'keto/ solves the phonological problem only at the considerably high price of resorting to a marked phonological rule, which has to be constrained so as to apply (i) to a specific grammatical form (the thematic vowel) (ii) in a specific morphological environment (the past participle slot) (iii) in all possible cases (some sort of absolute neutralisation relative to a specific paradigm slot, there being no surface true case of regular past participle in /eto/). Morphomic indexing is thus only apparently got around. In fact, it has to be invoked as a factor conditioning rule application. Again, a morphomic analysis of these exceptions makes them the paradigmatically-unmarked output, the weak ending *uto* being systematically co-selected with  $S_1$ , corresponding, for the subclass of verbs at hand, to the palatalised alternant.

#### 4. Morphomic productivity

There are two important consequences following from the type of evidence illustrated in section 3. Firstly and most apparently, ODS plays a prominent role as an active grammatical constraint on the morphological output of a native speaker of Italian. This looks like an across-the-board factor: when a specific stem alternant is assigned to a partition class, all slots belonging to the partition class in question are required to host that stem alternant, independently of phonological considerations. Phonological triggers are overridden by ODS. This is true independently of the classification of partition classes in section 3. Synchronically, the historical motivation of a partition class is relatively immaterial. It does not matter, for example, if  $S_2$  was originally based on a uniform phonological characterisation of its typical embedding phonological context (such as 'stressed stem AND ensuing back vowel'). Synchronically,  $S_2$  appears to be entirely crystallised into a purely morphomic category. Changes in the phonological context do not seem to synchronically affect the phonological identity of  $S_2$  in the least.

Still the neat characterisation of  $S_2$  in terms of phonological categories suggests that  $S_2$  might have had a phonological origin. Paradigmatic partition classes could then be looked at as remnants of originally phonologically-motivated processes, which eventually lost their phonological transparency and were reinterpreted morphomically (paradigmatically). As observed by Kiparsky (1982), this may be due to either of the following situations:



- (5) A rule  $A \rightarrow B / C\_D$   
 is opaque to the extent that there are surface representations of the form
- (i) A in environment C\_D, or
  - (ii) B in environment other than C\_D

In some respects, crystallisation of a partition class as morphomic is reminiscent of the well-known process of morphologisation of sound change (as in English *mouse/mice*). The main difference here is that morphomic crystallisation does not appear to correlate with morphological meaning in any straightforward way, if we exclude its partial correlation with **morphosemantic markedness**, as illustrated for Italian by Matthews (1981; 1991). This may have consequences on the way the relationship between—say— $S_1$  and  $S_2$  is perceived, although this hypothesis will not be explored further in this paper.

It is important to point out at this stage that, if we compound these remarks with two of the criterial properties identified by Aronoff for basic stem types, namely (i) stem types have no meaning, and (ii) they distribute predictably in language, then we come very close to identifying a distinctive notion of (relative) **morphomic productivity**. This notion does not seem to comply with general assumptions concerning language compositionality according to which the interpretation of a word form is a function of the interpretation of its constituent parts (and thus also of its stem). Secondly, unity of meaning and form is taken care of only at the level of the fully inflected form in the relevant paradigmatic slot (or the sign level, according to Beard's (1995) reuse of the Saussurean notion of sign); basically, all other intermediate stages of morphological realisation follow purely morphological constraints.

### 5. Paradigmatically-constrained phonology

The second main consequence of the state of affairs sketched in section 3 is that minor phonological rules appear to undergenerate and overgenerate hopelessly. This makes it very hard to state them formally in terms of paradigmatically-constrained phonological rules, as tentatively suggested by several authors in the early 70's (Kiparsky 1982; Harris 1973; Wanner 1972), but never worked out in practice, due to the persisting vagueness of notions such as 'paradigm' and 'stem allomorphy'. We can illustrate this point in some detail through the technical apparatus introduced in these pages.

The Italian verb CRESCERE 'grow' undergoes velar palatalization in all front vowel environments, with the exception of the past participle slot, where

palatalization takes place in what is an overtly bleeding phonological environment: /kre'f:uto/ instead of expected \*/'kres'kuto/. Moreover, CRESCERE exhibits a further stem alternant (/kreb:/) in some slots of the past absolute (S<sub>5</sub> in ODS). We assume that, after application of some morpholexical rules and indexing schemata, the morpho-phonological portion of the lexical entry of CRESCERE would look like this:

$$(6) \quad /kre \left[ \begin{array}{l} b: \quad /S_5 \\ sk \\ +VP \quad /otherwise \end{array} \right] /$$

The entry contains two stem alternants of CRESCERE and indexes them appropriately. As to the remaining third alternant, we can further assume that a velar palatalization rule<sup>7</sup> is at work, stated as follows:

(7) VELAR PALATALIZATION

$$\left[ \begin{array}{l} k \\ g \end{array} \right] \rightarrow \left[ \begin{array}{l} tʃ \\ dʒ \end{array} \right] / [ \text{---} [i, e, u, j] -s_2 ]$$

In this example, velar palatalization (VP) has the status of a minor rule. Hence it applies to those entries only which are specified as '+VP'. The rule is required to interact with ODS in the following way: it applies in all possible triggering environments, provided that they are not found in a slot marked as S<sub>2</sub>. This is conveyed by the '-S<sub>2</sub>' diacritic in the context part of the rule. This is useful to account for cases of underapplication of palatalization in the present subjunctive slots, as in erroneous overregularizations such as the form ?/'kreskino/ '(that) we grow'.

Injection of a paradigmatic diacritic in the context part of Velar Palatalization rule is not very enlightening, as it does not make any direct contact with ODS. For example, it requires that the same '-S<sub>2</sub>' be marked in the context part of any other minor phonological rule facing the same inconsistency (e.g., diphthongization of mid vowels in open syllables, compare the overregularization ?/'vengino/ with unattested \*/'vjenino/). In fact, the use of paradigmatic diacritics in the body of a phonological rule could be avoided by resorting to lexical indexing and a general blocking condition on rule application. Suppose that the phonological portion of the entry of CRESCERE is as follows:

<sup>7</sup> The rule accounts for only part of the derivation: namely *cres*[k] → *cres*[tʃ]. For simplicity, the other necessary steps are not considered, as they are immaterial to the argument developed here.

$$(8) \quad /kre \left[ \begin{array}{l} b: /S_5 \\ sk / S_2 \\ +VP / otherwise \end{array} \right] /$$

The stem /kresk/ is assigned the index  $S_2$  explicitly. Again, 'otherwise' signifies that it can also act as a default stem. On the other hand, /kreb:/ is assigned  $S_5$ , the partition class of past absolute. We can now formulate the following principle of paradigmatic lexical blocking of rule application:

- (9) If a stem  $S$  in a lexical entry is explicitly associated with a certain index  $I$ , no minor phonological rule can apply to  $S$  in the paradigmatic slots which are marked by  $I$  in ODS

Given this principle and the revised lexical entry of CRESCERE, /kresk/ would skip velar palatalization in all slots marked in ODS as  $S_2$ .

It is important to emphasise at this point that addition of /u/ in the triggering environment for velar palatalization accounts for the attested form /kre'f:uto/ 'grown' in the past participle slot. The move is admittedly unnatural phonologically. Alternatively, one can assume an underlying \*/kres'keto/, with successive application of palatalization and conversion of the thematic vowel into /u/. This solution, as already pointed out above, strikes us as equally unmotivated. However, there do not seem to be any other viable alternatives. Overapplication of phonological processes induced by ODS poses in fact a serious challenge to the phonological treatment of cases of paradigmatic levelling involving a default stem which does not coincide with the alleged phonological underlying form. The moral of this story is that paradigmatic conditioning of phonological processes goes a long way towards capturing aspects of the interaction between ODS and phonological reality. Still it fails to provide an entirely satisfactory account of all data. We suggest that the interaction between phonological and morphomic constraints should be looked for elsewhere. Another alternative, known as the **morpholexical approach**, has repeatedly been advocated in the literature: store stem alternants in the lexicon. It will be considered in the coming section.

## 6. The morpholexical solution

The sort of descriptive inadequacy of phonological accounts of stem alternation is not incurred by morpholexical treatments (Hudson 1974; Lieber 1981; Spencer 1988). In effect, what one seems to require for the data presented here to be appropriately described is just to split the generative rewriting rule

into two parts: the phonological change induced by the rule (accounting for stem alternation), and the context of its application (accounting for stem distribution). Morpholexical rules of some kind could well serve the purpose of capturing formal redundancies in stem alternation, by deriving all predictable alternants from a possibly underspecified base. On the other hand, stem indexing schemata (see Table 1 above) can be used to capture the contextual dimension of the distribution of the stem alternants thus obtained.

Although it is certainly true that the morpholexical solution cursorily sketched here is weakly equivalent to the generative rewriting model, it is not obvious that the former could also match the latter's explanatory power. At the root of the classical generative account lies the assumption that phonological changes are induced by context, and this is credited for explaining why stems exhibit a specific, phonologically constrained range of alternations. Morpholexical rules do not seem to be equally constrained. Furthermore, there are some aspects of ODS which are not immediately captured by the division of labour between morpholexical rules and stem indexing. To illustrate, let us consider a concrete example. When the partition class associated with  $S_5$  (covering some past absolute forms) is assigned to a stem alternant different from  $S_1$ , then a specific series of (unstressed) past absolute terminations must be selected: /i/, /e/, /ero/. They are traditionally called strong terminations, as opposed to default /'ei/, /'ette/, /'e/, /'erono/, /'ettero/ (also referred to as weak terminations). On the other hand, if  $S_1 = S_5$ , then the weak terminations are to follow. We have here a situation where one stem alternant ( $S_5$ ) correlates with a particular series of strong endings (see *crebbi* : *crebbe* : *crebbero*) if and only if  $S_5$  does not coincide with  $S_1$ . This type of co-selection is highly systematic: the alternation between  $S_1$  and  $S_5$  in the past absolute paradigm is directly governed by position of stress.  $S_5$  appears when stress falls on the stem;  $S_1$  does when stress falls on the ending. However, there is no general way to state this generalisation **directly** in the lexicon, if the phonological relation between  $S_1$  and  $S_5$  (accounting for stem alternation) and the context of its application (accounting for stem distribution) are dealt with separately. Observe that stem indexing alone is of no avail here. With regular verbs, it is the default alternant  $S_1$  which takes over the partition class of  $S_5$ . In this case, however, identity between  $S_5$  and  $S_1$  requires selection of weak terminations. Kiparsky (1996) suggests tackling these cases through full lexical listing of stem alternants, which are nonetheless selected **phonologically** on the basis on their stress patterns. Kiparsky's solution presents some technical difficulties, and, in fact, does not dispense with stem indexing. A full discussion of this point would lead us too far. Suffice it to say here that, once more,

the evidence considered here bears witness to the intricacy of the relationship between phonological and morphomic constraints, while leaving open the issue of where this tight relationship originates from.

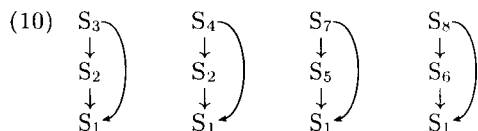
To sum up, the morpholexical approach to morphomically-governed stem alternation proves to be preferable over phonological accounts, however conditioned. While the latter are simply inadequate, the former is descriptively adequate, in spite of its failure to conveniently capture some relations between ODS and phonological principles in the grammar.

## 7. Historical language change

What can partition classes tell us about historical phenomena such as regularisation (levelling) and extension (or polarisation)? Due to levelling, existing stem alternations can be either curtailed or eliminated altogether, with the result that allomorphs of some morphemes become more similar to each other or merge completely. Thus the change of *horatos* : *horatoris* to *horator* : *horatoris* is a levelling of the alternation *s-r* (rhotacism) in stem final position. A typical example in Italian conjugation is formation of *seppellito* 'buried' on the basis of  $S_1$  (plus weak past participle ending), alongside with the strong participle *sepolto*.

The notion of re-indexing is particularly useful in this context. First, it suggests that partition classes always reduce in number through simple set union (in accordance with the Schema Transition Hypothesis, hereafter STH). Note that theoretical accounts which explain levelling as a mere process of global simplification of an historically earlier stage of grammar simply fail to predict that levelling takes place through simple set union of existing partition classes.

It is useful to represent the whole range of attested ways in which a stem variable is re-indexed in Italian conjugation as shown in (10) below, where re-indexing is pictorially represented by an arrow. Note that, due to STH, re-indexing maps one stem variable into one and only one other stem variable. If more than one arrow leaves the same stem index in (10) below (as, for example, in the case of  $S_4$ ) this means that the stem variable in question can be re-indexed in two alternative ways, thus giving rise to two distinct distribution schemata.



The diagram in (10) highlights the role of  $S_1$  as a kind of ultimate default for all stem alternants in Italian conjugation. Note that there also exist some secondary or intermediate defaults, corresponding to indices  $S_2$ ,  $S_5$  and  $S_6$ , which can be both the output and the input of re-indexing. They represent intermediate defaults in the sense that they are themselves ultimately merged into  $S_1$  through a further step of morphomic regularisation.

Extension, on the other hand, is more difficult to model through re-indexing. Extension refers to a type of analogical change in which existing alternations spread to new instances. An interesting example in Old Italian is the etymological alternation *giungo* : *giugniamo* which produced, according to Rohlfs (1966), non-etymological *vengo* : *vegiamo* (from earlier *vegno* : *vegiamo*) by ‘analogy’. Note that, here, linguistic contrasts come to be more fully implemented than before, whereas levelling has precisely the opposite effect. In our terms, a further partition was introduced in the paradigm of a class of verbs where there was one partition only. How can we account for this? If, as suggested above, phonological change has a bearing on the establishment of partition classes (through a morphomic reanalysis of earlier phonologically-motivated contexts), it is reasonable to expect that it should also play a role in the way re-indexing takes place historically. Loss of transparency of the phonological rule that produced *vegno* from Latin *veneō* might have resulted in perceiving the ending *ngo/gno* in the first person singular of present indicative as two stylistically-related variants. Eventually, the normative pressure of the Latin etymological pronunciation may have led to an extension of *ngo* in more and more assorted contexts. If this hypothesis is plausible, we should then expect to spot traces of the same sort of phonological reanalysis in contemporary Italian.

Italian conjugation offers at least one interesting case of ongoing levelling. The verb *SEDERE* ‘sit’ exhibits a classical example of stem alternation due to diphthongization of mid vowel in open stressed syllables (see Table 7 below).

Table 7

| PRESENT INDICATIVE |         |         |            |           |           |
|--------------------|---------|---------|------------|-----------|-----------|
| SINGULAR           |         |         | PLURAL     |           |           |
| 1                  | 2       | 3       | 1          | 2         | 3         |
| /sʝedo/            | /sʝedi/ | /sʝede/ | /se'djamo/ | /se'dete/ | /sʝedono/ |

There is a tendency, in contemporary Italian, towards analogical extension of diphthongized stem alternants in phonologically unmotivated contexts. To wit, forms such as /mwo'vete/ and /skwo'tete/ have gradually replaced etymological /mo'vete/ and /sko'tete/, which are still in marginal use as regional variants. Another interesting case of this extension is illustrated by the future indicative and present conditional forms of SEDERE, showing a marked tendency towards extension of the diphthongized stem in all relevant slots (/sʝede'rete/, /sʝede'reste/). Note that here levelling does not go in the direction of the default stem. Why then a tendency towards the diphthong?

The first thing to be noted here is that diphthongization of mid vowels in stressed open syllables is no longer an automatic process in the phonology of contemporary Italian. Secondly, a native speaker of Italian is likely to feel a still prominent phonological correlation between a monophthong and its corresponding diphthong. This is mainly due to the persistence in regional/dialectal areas of monophthongized variants of what is normally pronounced with a diphthong in standard Italian (/bɔno/ : /bwɔno/, /so'nate/ : /swo'nate/ etc.). It is then by no means implausible that a directional phonological derivation, from the monophthong to the diphthong, is eventually replaced by a bidirectional relation, whereby the two alternants are perceived as mutually related, context-free variants. For lack of a better term, we would like to dub this sort of phonological correlation as paradigmatic.<sup>8</sup> We can then suppose that cases of phonological correlation based on this type of evidence can give rise to instances of levelling that go in either direction, both compatible with STH. Preference for the diphthong is then justified on the basis of a stylistic choice, the monophthongized stem showing a distinct vernacular connotation, while the diphthong

<sup>8</sup> This analysis differs from that proposed in Andersen (1972), where the direction of original phonological rules is abductively reversed through **adaptive rules** of some kind. In our proposal, the phenomenon is only partly explained in grammatical terms, the direction of levelling being dictated by preferences which are stylistic rather than grammatical in a strict sense. This is confirmed by the fact that one finds, in Old Italian, examples of extension towards the opposite end, e.g., *trovo* from \**truovo*, or *piagno* from \**piango* (Rohlf's 1966).

being perceived as more standard. At present, we can only speculate about the reasons why  $S_6$  is the starting point of this instance of levelling in progress. Notions such as low salience of the partition class  $S_6$  (Kiparsky 1982) can conspire with the phonological similarity between monophthong and diphthong to make this type of merger more likely than other equally conceivable ones.

## 8. Paradigms and language learning

So far, we have mainly been entertaining the hypothesis that the interaction of paradigmatic and syntagmatic principles of stem alternation should be accounted for as a purely grammatical phenomenon **within** the linguistic system. This hypothesis dates back to the early 70's. We showed that it can be refined considerably by making use of notions such as stem indexing, paradigmatic partition classes and the like. Still the resulting account is not entirely satisfactory. Historical evidence too is only partly in line with formal descriptive devices such as ODS and general principles such as STH. This prompts the suggestion that the range of phenomena which have been the focus of the present paper actually originates from the interaction of the grammatical system with other more use-oriented systems, such as language learning, production and perception. In this section, we would like to picture a scenario involving a particular family of language learning algorithms, as developed in Artificial Intelligence circles.

Of late, considerable interest has been devoted to **memory-based** machine language learning, a family of self-learning algorithms which all share the property of grounding generalizations about unknown linguistic objects on their more or less direct analogy to already known examples. This approach is also known as **lazy learning**, since the amount of actual learning involved in it is comparatively negligible, and mainly consists in rote memorization of examples (Daelemans 1996; Mitchell 1997). Here, we would like to bring to the reader's attention one particular member of this family, paradigm-based analogical learning (Pirrelli-Federici 1993; 1994; Federici-Pirrelli-Yvon 1996; Pirrelli-Yvon 1998; Lepage-Shin-Ici 1996; Lepage 1998), which makes use of the notion of proportional analogy in a way which seems to us particularly relevant to the contents of the present discussion.



Paradigm-based analogical learning is concerned with analogies that are based on an underlying systematic relation of proportionality.<sup>9</sup> A proportion is not established by an existing analogy; rather it is a proportional relationship that establishes the analogy. Accordingly, analogy is not an inherent relationship between any two terms, but a recurrent proportionality between two series of terms. The following discussion is considerably indebted to Pirrelli-Yvon (1998; 1999).

There are several different ways of constructing a proportion to establish an analogy. For example, consider the following morphological proportion:

$$(11) \quad \begin{array}{cccc} \text{A} & \text{B} & \text{C} & \text{D} \\ /'wɔ:k/ & : & /'wɔ:kt/ & = & \text{WALK/Verb/PRESENT} & : & \text{WALK/Verb/PAST} \end{array}$$

where the first two terms of the proportion (A and B) are word forms, and the last two (C and D) are bundles of morpholexical tags. Proportion (11) establishes a correspondence between units defined at different levels of linguistic representation (what is shared by A and B should be paired with what is common between C and D), and amounts to bootstrapping a lexicon of morphemes from known word pairs. A principled difficulty with this approach is widespread lack of isomorphism between minimal units of content and form in the morphology of inflecting languages. Problems of this kind are well known in the linguistic literature (Matthews 1972; 1991), and led Pike (1963), among other morphologists, to cast doubts on the issue of word-internal segmentation. More recently, Beard (1995) has considerably clarified the issue, by suggesting that the segmentation fallacy in morphology be tackled through the "separation hypothesis" (see also Hudson 1984). Beard resumes the structuralist notion of the word as a sign, that is, a complex structure of units defined over three levels of linguistic analysis: phonological, grammatical and semantic. Linguistic processes are supposed to operate at any such level on the basis of autonomous principles. This does not mean that there exists no correlation between any two such levels. Beard argues that this correlation is **indirect** (governed by the operation of autonomous principles) and **sign-based**, since it is the notion of the word as a sign which ultimately guarantees that operations defined at different levels are eventually kept in step.

<sup>9</sup> This qualification clears the ground of other conceivable uses of the word analogy, e.g., as some sort of direct, pairwise similarity as suggested by Skousen (1989). Paradigm-based Analogy shares some assumptions with the Network Model (Bybee 1988), although lack of running implementations of Bybee's model makes it impossible to make a detailed comparison.

This idea can be expressed proportionally by assuming that the terms of an analogical proportion should be linguistic signs in their own right. A well-defined linguistic proportion consists of a system of interlocked level-wise proportions, each holding at one linguistic level only. The overall sign-based proportion obtains if and only if proportionality holds **simultaneously at all levels involved**. This ensures that operations autonomously defined at each level are nonetheless correlated. Accordingly, the proportion in (11) above, which directly associates units pertaining to different domains, should be restated as the following, more complex system of two proportions (a phonological and a morpholexical one), obtaining simultaneously:

$$(12) \quad /wɔ:k/ : /wɔ:kt/ = /tɔ:k/ : /tɔ:kt/ \\ \text{WALK/Verb/PRES} : \text{WALK/Verb/PAST} = \text{TALK/Verb/PRES} : \text{TALK/Verb/PAST}$$

Unlike proportion (11) above, sign-based proportions like (12) no longer take care of the correspondence between units defined at different levels. From the perspective of language learning, proportions such as (11) support the view that language acquisition and language processing boil down to the establishment of a possibly isomorphic correspondence between minimal linguistic units of different domains (levels). To the contrary, use of multi-level proportions such as (12) reflects a view of language learning and processing as a sign-based activity, whereby linguistic signs are mutually related through a “system” of proportions, one for each relevant linguistic level. Pirrelli and Yvon (1998) use analogical proportions such as (12) to infer, e.g., that the unknown form /tɔ:kt/ is the past tense of TALK, on the basis of the proportionality holding among the three other terms. Since other proportions may lend support to different analyses, Pirrelli and Yvon suggest preferring the analysis supported by the maximum number of proportions.

However comparatively simple, this operational notion of proportional analogy is effective in modelling the way ODS tightly constrains the logical problem of learning inflection. Observe first that partition classes in ODS are complied with by both regular and subregular verbs. If alternative proportional analyses are available, a majority vote, in terms of number of supporting paradigms, will always favour analogies involving forms which belong to the same partition class, since, as we saw, **they virtually obtain in all cases**, no matter how regular the verbs in question are. On the contrary, for what we showed so far, analogies of forms belonging to different partition classes will hold only for a subset of verbs (mainly regular ones).

This model has the potential of shedding light on two important findings: (i) isomorphism of form and meaning, although a semiotically desirable

property (Dressler 1987), does not seem to play a central role in morphological learning (Beard 1995), and (ii) analogical learning tends to put a premium on minimal extensions which obtain for the maximum number of verbs, rather than privilege sweeping generalisations (involving a maximum number of paradigm slots) which however hold for a subset of verbs only. This suggests that the analogical learning of morphology by a child has, as its ultimate target, a “meta-paradigm” such as the one in Table 4, rather than a sheer collection of relatively unrelated paradigms. In other words, learning tends to favour inter-paradigmatic consistency over intra-paradigmatic uniformity. If these findings are confirmed, ODS can ultimately be explained as the by-product of a more basic function of analogy: namely its key role in language acquisition. Finally, both findings are in keeping with the notion of morphomic productivity outlined in this paper.

## 9. Conclusion

We discussed some formal consequences of Aronoff’s (1994) notion of basic stem type, in the light of data from the Italian conjugation. We pointed out that morphomic partition classes stake out a fruitful dimension of linguistic analysis, which can naturally accommodate insights concerning conspiracy of constraints on morphonological processes. This analysis led us to two non-trivial conclusions: (i) minor phonological rules exhibit a morphomic behaviour synchronically, (ii) morphomic partition classes betray a phonological origin historically.

Furthermore, we delved into the notion of morphomic productivity, a purely paradigmatic mechanism of word formation, which *prima facie* departs from traditional assumptions of strict compositionality. This notion appears to be justified on descriptive grounds.

In an attempt to support descriptive adequacy with a more explanatory account, we tried to establish a connection between morphomic productivity, language change and language learning, through a unifying approach to machine language learning based on paradigm-driven analogy. This move reflects the intuition that the range of phenomena which have been the focus of the present paper actually originates from the interaction of the grammatical system with other more use-oriented systems, such as language learning, production and perception. We conclude that: (a) recent findings in the micro-structure of paradigms have considerably sharpened our appreciation and understanding of the role of paradigms in the grammar system of human speakers;

(b) as this paper and other related papers have shown, we hope, extensively, there are compelling arguments in support of the view that the morphomic (or paradigmatic) structure of the conjugation system of a language plays a prominent role in selecting the synchronically correct grammar, independently of historical considerations; (c) this view is bound to lead to a better understanding of phenomena of language change; (d) last but not least, this perspective calls for a tighter co-operation between Linguistic Theory, models of Natural Language Processing and Machine Language Learning, along the lines sketched in Bybee (1996).

### References

- Andersen, H. 1972. Abductive and deductive change. In: *Language* 49: 765–93.
- Aronoff, M. 1994. *Morphology by itself*. MIT Press, Cambridge MA.
- Beard, R. 1995. *Lexeme-morpheme base morphology: a general theory of inflection and word formation*. State University of New York Press, Albany.
- Bybee, J. 1988. Morphology as lexical organization. In: Hammond, M. – Noonan, N. (eds): *Theoretical morphology. Approaches in modern linguistics*, 119–41. Academic Press, San Diego.
- Bybee, J. 1996. Productivity, regularity and fusion: how language use affects the lexicon. In: Singh, R. (ed.): *Trubetzkoy's orphan: proceedings of the Montreal roundtable "Morphology: contemporary responses"*, *Current Issues in Linguistic Theory*, 144, 247–79. John Benjamins, Amsterdam and Philadelphia.
- Carstairs, A. 1990. Phonologically conditioned suppletion. In: Dressler, W.U. – Luschützky, H.C. – Pfeiffer, O.E. – Rennison, J.R. (eds): *Contemporary morphology*, 17–23. de Gruyter, Berlin.
- Daelemans, W. 1996. Abstraction considered harmful: lazy learning of language processing. In: van den Herik, H. – Weijters, A. (eds): *Proceedings of the 6th Belgian-Dutch Conference on Machine Learning*, 3–12. Maastricht.
- Dressler, W.U. 1987. Word formation as part of natural morphology. In: Dressler, W. – Mayerthaler, W. – Panagl, O. – Wurzel, W. (eds): *Leitmotifs in Natural Morphology*, 99–124. John Benjamins, Amsterdam and Philadelphia.
- Federici, S. – Pirrelli, V. – Yvon, F. 1996. A dynamic approach to paradigm-driven analogy. In: Wermter, S. – Riloff, E. – Scheler, G. (eds): *Symbolic, connectionist and statistical approaches to natural language processing*, 385–98. Springer-Verlag, Berlin.
- Harris, J.W. 1973. On the order of certain phonological rules in Spanish. In: Anderson, S. – Kiparsky, P. (eds): *A Festschrift for Morris Halle*, 59–76. Holt, Rinehart and Winston, New York.
- Hudson, G. 1974. The representation of non-productive alternation. In: Anderson, J. – Jones, C. (eds): *Historical linguistics*, 203–29. North Holland, Amsterdam.

- Hudson, R. 1984. *Word grammar*. Blackwell, Oxford and Cambridge MA.
- Kiparsky, P. 1982. *Explanation in phonology*. Foris, Dordrecht.
- Kiparsky, P. 1996. Allomorphy and morphophonology. In: Singh, R. (ed.): *Trubetzkoy's orphan: proceedings of the Montreal roundtable "Morphonology: contemporary responses"*, *Current Issues in Linguistic Theory*, 144, 13–63. John Benjamins, Amsterdam and Philadelphia.
- Lepage, Y. 1998. Solving analogies of words: an algorithm. In: *Proceedings of COLING-ACL '98*, 728–35. Montreal.
- Lepage, Y. – Shin-Ici, A. 1996. Saussurean analogy: a theoretical account and its application. In: *Proceedings of COLING-96*, 717–22. Copenhagen.
- Lieber, R. 1981. *On the organization of the lexicon*. Indiana University Linguistic Club, Bloomington.
- Lightner, T. 1968. On the use of minor rules in Russian phonology. In: *Journal of Linguistics* 4: 69–72.
- Matthews, P.H. 1972. *Inflectional morphology. A theoretical study based on aspects of Latin verb conjugation*. Cambridge University Press, Cambridge.
- Matthews, P.H. 1981. Present stem alternation in Italian. In: Geckeler, H. – Schlieben-Lange, B. – Trabandt, J. – Weydt, H. (eds): *Logos semanticos: studia linguistica in honorem Eugenio Coseriu*, 57–65. de Gruyter, Berlin.
- Matthews, P.H. 1991. *Morphology (Second edition)*. Cambridge University Press, Cambridge.
- Mitchell, T. 1997. *Machine learning*. McGraw-Hill, Boston MA.
- Pike, K.L. 1963. Theoretical implications of matrix permutation in Fore (New Guinea). In: *Anthropological Linguistics* 5.
- Pirrelli, V. – Battista, M. 1998. *The paradigmatic dimension of stem allomorphy in Italian inflection*, ILC-CNR Technical Report. Pisa, Italy.
- Pirrelli, V. – Federici, S. 1993. An analogical way to language modelling: MORPHEUS. *Acta Linguistica Hungarica* 41: 235–63.
- Pirrelli, V. – Federici, S. 1994. 'Derivational' paradigms in morphonology. In: *Proceedings of COLING-94*, Kyoto.
- Pirrelli, V. – Yvon, F. 1998. Analogy in the lexicon: a probe into analogy-based machine learning of language. In: *Proceedings of the 6th International Symposium on Human Communication*, 576–81. Santiago de Cuba.
- Pirrelli, V. – Yvon, F. 1999. The hidden dimension: a paradigmatic view of data-driven NLP. In: *Journal of Experimental and Theoretical Artificial Intelligence* 11: 391–408.
- Rohlf, G. 1966. *Grammatica Storica della Lingua Italiana e dei suoi Dialetti*. Einaudi, Torino.
- Scalise, S. 1981. Il suffisso *Xione* e alcune ipotesi della morfologia lessicale. In: *Rivista di Grammatica Generativa* 6: 155–79.

- Scalise, S. 1983. *Morfologia lessicale*. CLESP, Padua.
- Skousen, R. 1989. *Arguments for modeling of language*. Kluwer, Dordrecht.
- Spencer, A. 1988. Arguments for morpholexical rules. In: *Linguistics* 24: 1–29.
- Stump, G.T. 1995. Stem formation, stem indexing, and stem choice. Ms.
- Wanner, D. 1972. The derivation of inflectional paradigms in Italian. In: Casagrande, J. – Saciuk, B. (eds): *Generative studies in Romance languages*, 293–318. Newbury House, Rowly MA.
- Zwicky, A. 1990. Inflectional morphology as a (sub)component of grammar. In: Dressler, W.U. – Luschützky, H.C. – Pfeiffer, O.E. – Rennison, J.R. (eds): *Contemporary morphology*, 217–36. de Gruyter, Berlin.

Address of the authors: Vito Pirrelli and Marco Battista  
Istituto di Linguistica Computazionale - CNR  
Area della Ricerca di Pisa  
via Alfieri 1  
56010, S. Cataldo  
Pisa, Italy  
vito@ilc.pi.cnr.it  
marco@ilc.pi.cnr.it

## DEVERBAL ADJECTIVALIZATION AS A GRADUAL PROCESS

PETRA SLEEMAN AND ELS VERHEUGD

### Abstract

In this paper, we defend the idea that, in languages such as English, German and Dutch, there are in principle three types of (de)verbal modifiers of the noun: the first, postnominal, type is a reduced relative clause, that is, a full verbal projection with an operator-like subject. Prenominal (de)verbal modifiers with an eventive meaning represent the second type. We claim that they are poorly inflected verbs with a reduced argument structure. The third type is the participle lexicalized as an adjective, which denotes a pure property and has completely lost the argument structure of the verbal base.

### Introduction

Participles are verbal forms that in different types of languages, such as Germanic and Romance, can be used as modifiers of the noun. Both for present and for past participles there has been an extensive discussion in the literature on their categorial status: have they become adjectives or are they still verbs in their use as modifiers (see e.g., Levin–Rappaport 1984 and Borer 1990)?

Several linguists argue that both for present and for past participles a distinction has to be made between adjectival and verbal forms. According to e.g., Bennis–Wehrmann (1990), present participles can be lexicalized as adjectives (see also Brekke 1988). This happens in the case of participial forms of psych verbs such as *astonishing* or *deceiving*. Since these forms can be modified by the adverb *very*, they must be adjectival:

- (1) a very astonishing remark
- (2) this method is very deceiving

Past participles that can be preceded by *very* are also generally assumed to be adjectival (Siegel 1973; Williams 1981):

- (3) a very pleased person

Hoekstra (1984) and Bennis–Wehrmann (1990) show that the same holds for Dutch present and past participles (*heel* is the Dutch equivalent of *very*):

- (4) een heel vervelend boek  
‘a very boring book’
- (5) een heel verbaasde man  
‘a very astonished man’

But whereas participial forms that can be preceded by *very* or *heel* clearly are adjectives, other participles that can be used as modifiers of the noun retain some of their verbal properties. The participles in the following Dutch examples can be combined with an object, a verb-modifying adverbial phrase or a *by*-phrase:

- (6) de aan zijn principes vasthoudende man  
‘the man holding on to his principles’
- (7) het in Hongarije gekochte boek  
‘the book bought in Hungary’
- (8) het door jou vertelde verhaal  
‘the story told by you’

As the examples show, these participles are in prenominal position in Dutch, just like simple adjectives (*het dikke boek* ‘the thick book’). In English, however, where simple adjectives precede the noun (*a long story*), they are postnominal, as the translations show. Whereas for English it is generally assumed that the postnominal participles are verbal (see e.g., Fabb 1984), for languages like German and Dutch it has been argued that the not lexicalized, more verbal participles belong to a neutralized [+V] category (see e.g., Hoekstra 1984; Toman 1986) or a mixed V/A category (Bennis–Wehrmann 1990; Haspelmath 1996). Van Riemsdijk (1981) and Fanselow (1986), however, claim for very similar German cases that they are verbs projecting clausal structure.

In this paper we argue that the type of participle exemplified in (6)–(8) is a verb and thus [+V, –N], despite its prenominal position and despite its adjectival inflection. We will not claim, however, that the participle projects clausal



structure and takes a (phonologically unrealized) syntactic subject (PRO according to van Riemsdijk and Toman, an operator-bound empty NP according to Fanselow). Our claim is that prenominal modifiers with verbal properties do not have a syntactic external argument and do not project clausal structure, but are related to the noun in a direct way, i.e., via theta-identification (Higginbotham 1985), just like adjectives, but without having the categorial status of an adjective. In our view deverbal adjectivalization is a gradual process that involves a gradual loss of syntactic argument structure and a gradual transition w.r.t. the meaning from an event denoting expression into a property denoting expression. We propose that participial modifiers of the noun can be divided into three types: those that have their complete argument structure and event structure, just like their verbal base, those that have lost their external or 'to be externalized' argument but are still eventive, and those that have lost their argument and event structure altogether. Only for this last category, exemplified in (1)–(5), will we claim an adjectival status. We show furthermore that the three types are not equally represented in the languages that we will discuss in this paper (English, Dutch and German). We will compare deverbal adjectivalization to other deverbal processes and we will show that they all involve a gradual loss of argument and event structure.

This paper is organized as follows. In section 1, we discuss the case of deverbal nominalization, which has also been argued to be a gradual process. In section 2, we present the data, which show that different types of deverbal modifiers have to be distinguished. In 3, we discuss several analyses that have been proposed in the literature for the modifiers in question. In 4, we present our own analysis. In section 5, we compare inflectional and derivational deverbal adjectivalization and nominalization. Finally, in 6, the results are summarized.

## 1. Nominalization as a gradual process

Van Hout (1991) investigates the morphological operation of nominalizing a verb. This operation is generally assumed to result in two types of nominal: the first type denotes an event (complex event or process nominal), the second type an object (result nominal). This semantic distinction has been correlated (see, for example, Grimshaw 1990) with the property of taking, respectively not taking syntactic arguments. The examples (9) and (10) illustrate the two types (note that the *of*-phrase in (10) is assumed to be an optional adjunct, not a syntactic argument):

- (9) The construction of that building took place in 9 months.
- (10) I do not like that translation of 'War and Peace'.

Now, Van Hout shows that this distinction is not fine-grained enough and that the process of nominalization is a more gradual one. She proposes that, in a first step, nominalization highlights the **state** node of the original event structure of the verbal base. The resulting event nominal is not countable, takes obligatorily one or more internal arguments and can be followed by a *by*-phrase expressing the original Agent (*the destruction(\*s) of the city (by the enemies)*). In other words, this process noun denotes an event and has argument structure, although the (original) external argument is no longer obligatorily mapped onto a syntactic position. In the second step, the denotation is transformed from an event into an object, but without deletion of the original event structure of the verb. The result is a countable event nominal, which can optionally be followed by an internal argument and a *by*-phrase or genitive phrase (*I attended all performances (of/by Jouri Egorov) (of the Schumann program)*). Since the event structure of the verbal base is not deleted, realization of the original arguments of the verb by adjuncts is possible. The deletion of the event structure is the final step and gives rise to a result noun (*the/John's proposals*), that takes no arguments. Here, *John* is a free adjunct that can have all kinds of interpretation, including that of the Agent of the proposal.

So, the nominalization operation consists in a gradual loss of the internal syntactic and logical properties of the original verbal base. Although the resulting form has the categorial status of a noun from the first step on, it becomes more and more 'nouny' during the process.

## 2. Different types of (de)verbal modifiers

As we saw in the introduction, at least two types of participial modifiers can be distinguished: the participle lexicalized as an adjective, and the participle with verbal properties. There is reason to believe that within the latter category two subcases have to be distinguished.

In English, modifiers sometimes occur in prenominal, sometimes in postnominal position. Simple adjectives differ from relative clauses w.r.t. their position:

- (11) a blue car

(12) the man that I saw in the park

But, as e.g., Sadler and Arnold (1994) observe, not only relatives are post-nominal: *to* + infinitive as well as present and past participles, which are all constructions based on a verb, are normally in this position:

(13) ***to* + infinitive**  
the book to read

(14) **present participle**  
the man sitting on the sidelines

(15) **past participle**  
the jewels stolen yesterday

In Sleeman–Verheugd (1998) we argue that these postnominal (de)verbal modifiers are reduced relative clauses with one of the arguments realized as an empty operator that moves to [Spec, CP]. This analysis is based not only on the postnominal position of the (de)verbal modifiers in question, but also on the fact that they can be combined with the pronoun *those*, just like relative clauses but unlike simple adjectives:

(16) those that I saw yesterday

(17) those to read

(18) those sitting on the sidelines

(19) those stolen yesterday

(20) \*those blue

As we showed in (6)–(8) above, Dutch (de)verbal modifiers occur to the left of the noun, just like adjectives but unlike relative clauses. The same is true for German (the examples containing participial forms have been taken from Toman 1986):<sup>1</sup>

<sup>1</sup> Whereas English uses *to* + infinitive, German uses *zu* + present participle. In section 5 we will discuss a construction in which the reverse holds: a present participle in English but an infinitive in German.

- (21) **zu + present participle**  
die von einem Arzt zu untersuchende Frau  
'the woman to be examined by a doctor'
- (22) **present participle**  
der seinen Vater rasierende Mann  
'the man shaving his father'
- (23) **past participle**  
der von allen gehasste Hausmeister  
'the janitor hated by everyone'
- (24) **simple adjective**  
ein schönes Kleid  
'a beautiful dress'
- (25) **relative clause**  
der Mann den du hasst ist der Hausmeister  
'the man that you hate is the janitor'

The Dutch and German facts raise the question of what the status of the prenominal (de)verbal modifiers, and in particular of the present and past participle, is: are they (reduced) clausal constituents, adjectives or simply verbs? In the next section we will discuss several positions that have been taken in the literature.

### 3. Previous analyses

There is general agreement in the literature that the present and past participles with a purely adjectival (that is, non-eventive), idiosyncratic meaning such as *astonishing* and *learned* in English, *ontroerend* 'touching' and *bemind* 'beloved' in Dutch and *überzeugend* 'convincing' and *geeignet* 'appropriate' in German are lexicalized deverbal adjectives (see, for example, Fabb 1984; Toman 1986; van der Putten 1997). An argument for this view is that in Dutch and German they participate in synthetic comparative formation, just like normal adjectives, as *ontroerender* 'more touching' and *überzeugender* 'more convincing' show.

Postnominal participles in English, which have always a verbal reading, are analysed by Fabb (1984) as clausal entities with a PRO subject and an empty Inflection, as in (26):

- (26) children [PRO INFL playing in the park]

In Sleeman–Verheugd (1998), we defend the idea that these postnominal participles are indeed clausal, but we propose an empty operator instead of PRO as the subject.

Fanselow (1986) also analyses German cases such as (22) above, with a prenominal present participle, as a reduced relative clause, with a phonologically empty NP bound by an empty operator in Comp as the phonetically unrealized argument:

- (27) der [Op<sub>i</sub> seinen Vater t<sub>i</sub> rasierende] Mann  
 'the man shaving his father'

The presence of an empty subject is motivated by binding facts as showing up in e.g., the German example (28): with a syntactic subject present, the anaphoric element can be bound within the modifier, so that no binding with the antecedent *Männer* (which would result in a violation of the **i-within-i Condition**, see Chomsky 1981) needs to be stipulated:

- (28) die [Op<sub>i</sub> einander t<sub>i</sub> verachtende] Männer  
 'the men despising each other'

A crucial point in Fanselow's analysis is that the empty, operator-bound NP, which bears a theta-role and is a variable in Logical Form, is identified via a general interpretational rule (namely lambda-abstraction) with the antecedent noun. The result is that the property denoted by the modifier and the property denoted by the noun combine into one property.

The main reasons for analysing prenominal present and past participles as some sort of verbal projections (in fact, a reduced relative is a full verbal projection) are that the forms in question have a verbal reading, take syntactic arguments (as in (27) and (28)), combine with time adverbials normally modifying events (e.g., *het gisteren gesloten zwembad* 'the swimming pool closed yesterday') and do not allow comparative formation (\**die mich noch enttäuscherenden Ergebnisse* 'the results disappointing me even more'). On the other hand they are like adjectives in being used as modifiers of the noun, and in taking adjectival inflection, as the German examples show. Because of this last point, prenominal participles have been analysed quite generally either as belonging to a mixed category V/A or to a neutralized category [+V].

Toman (1986) and van der Putten (1997), for example, argue for German, respectively Dutch, that *-end* is an adjectival affix taking VP as its sister. The operation in question is therefore not a lexical one, but takes place in syntax.

The same claim is made for past participles by van der Putten, with the difference that the adjectival affix adjoined to VP is taken to express perfective aspect. According to Toman, past participles are derived from the underlying verb by a word formation process that operates on the [N] feature of the base. The result is a neutralized category [+V]. Lexical insertion under non-distinctness would allow for the insertion of such a category into an AP position.

Wunderlich (1987), who takes a lexicalist point of view, considers German prenominal participles to be forms produced by the morphological component, and not in syntax. Both past and present participles are, in his view, optionally converted from a lexical item of the category [+V, -N] into a form with the feature specification [+V, +N]. Because of the optionality of this rule, both types of participles would be produced by the morphological component. In prenominal position, the [+N] participle would have to be used. In order to account for the fact that the theta-grid as well as the eventive meaning of the underlying verb remain in principle intact after conversion, Wunderlich assumes that the affixes used to form the participles in question are inflectional, and are therefore transparent. This does not mean, however, that the verb projects all its remaining arguments. Wunderlich clearly distinguishes between the semantic arguments of a verb (what one could also call its participants at the level of Lexical Conceptual Structure) and its syntactic arguments. When a lexical category has for example a (semantic) external argument, this argument can, but need not, be projected syntactically. This situation arises in the case of prenominal modifiers of the noun (e.g., *die kluge Frau* 'the intelligent woman', *der geschriebene Brief* 'the letter written', *die sich umarmende Frauen* 'the women embracing each other'). Although lexical entities such as the adjective *klug* and the noun *Frau* both take an external argument, logically speaking, they do not project it syntactically in *die kluge Jungfrau*. The two external arguments are assumed to unify at the level of Logical Form, by which process a conjunction of properties arises. No syntactic empty elements (such as PRO or an empty, operator-bound NP) need to be postulated. Binding phenomena as exemplified by (28) above can be dealt with within the semantic component. The unification of arguments proposed by Wunderlich is very close to the operation of theta-identification, taken from Higginbotham (1985), that we will propose for prenominal, deverbal modifiers.

Like Wunderlich, Haspelmath (1996) adheres to the Lexicalist Hypothesis and claims that participles are formed by an inflectional affix that nevertheless changes the word class. So, inflection can be category-changing, in his view. What Haspelmath tries to show is that the distinction between inflection and derivation is not an absolute, but a gradual one; it would correlate furthermore

with particular syntactic properties of the resulting forms, namely an equally gradual distinction between more, respectively less, preservation of the internal syntax of the base forms. The problem with participles, in this respect, is that they behave externally like adjectives, but internally like verbs. In other words, there seems to be a mismatch between morphology and syntax. In order to solve this problem, Haspelmath proposes to distinguish several word-classes within a word, that is, the word-class of the base lexeme, relevant for internal syntax, and the word-class of the inflectional transferring affix that determines the word-class of the whole word. The category of the whole word is relevant for external syntax.

In the next section we will claim that participles are indeed inflectional forms, without assuming, however, that there is a change in category. We will try to show furthermore that the problem of lexical insertion and that of adjectival agreement, which are, for most linguists, the main reasons for claiming an adjectival status for participles, can be dealt with within the proposed analysis.

#### 4. The proposed analysis

We have distinguished three kinds of deverbal modifiers: postnominal (de)verbal modifiers in English, prenominal (de)verbal modifiers in Dutch and German, and lexicalized participles.<sup>2</sup> We assume that postnominal (de)verbal modifiers such as participles in English are verbs projecting a complete argument structure.<sup>3,4</sup> The external or 'to be externalized' argument is realized as an empty operator; these modifiers are therefore (reduced) clausal entities expressing an event:

(29) the child [Op singing a song]

(30) the book [Op bought by John]

<sup>2</sup> The three types are not realized in all languages. Dutch and German do not have reduced relatives. For an explanation see below.

<sup>3</sup> We assume, following Grimshaw (1990), that *by*-phrases are the realization of an LCS-participant rather than a syntactic argument. In Grimshaw's view this means that there has been suppression of the argument from syntactic argument structure.

<sup>4</sup> Prepositional phrases that function as modifiers such as in *the man with the hat* are not taken into account in this paper. In Sleeman – Verheugd (1998) we analyse them as reduced relative clauses with an empty subject.

Lexicalized participles, in all three languages, are adjectives. They denote a pure property:

(31) a very boring film

(32) the very astonished man

In this paper we are mainly concerned with the analysis of prenominal 'verbal' participles in Dutch and German, which behave internally like verbs (so that they can assign Case, see the German example (22)), but externally like adjectives and which denote both an event and a property. Since we respect the Lexicalist Hypothesis, we reject the syntactic derivation solution as proposed by van der Putten (1997).<sup>5</sup> Since a double categorial status as proposed by e.g., Haspelmath (1996) or feature neutralization as proposed by Toman (1986) are not attractive from a theoretical point of view, we propose an alternative in this section.

What we would like to propose is that the adjectival behaviour of the prenominal (de)verbal modifier is the consequence of the impoverished syntactic argument structure of the participle, without there being a categorial change from a verb into an adjective. In our view, the syntactic and semantic differences between relative clauses, prenominal (de)verbal modifiers and (deverbal) adjectives is the consequence of a process of gradual reduction of syntactic argument structure and event structure. Whereas the verb in a relative clause has its complete argument and event structure, simple adjectives have no syntactic argument structure at all and no event structure, either.<sup>6</sup> (De)verbal modifiers in Dutch and German are in an intermediate position: they are eventive but have a reduced argument structure. What is missing from the argument structure is the external or 'to be externalized' argument. Since this argument is missing, realization as PRO (Fabb, Toman) or an operator-bound empty subject (Fanselow) is not possible. But although one of the arguments is syntactically not realized, semantically it is still there: the theta-role is present in the Lexical Conceptual Structure (LCS) of the verb, but is not mapped onto a syntactic

<sup>5</sup> Although we respect the Lexicalist Hypothesis, we will assume below that non-verbal inflection on verbal forms is the result of a syntactic operation.

<sup>6</sup> Although, in our view, prenominal attributive adjectives are related to the noun in a direct way, i.e., without the projection of any argument, the projection of the theme is required if a clausal structure has to be formed, i.e., if the adjective has to function as the predicate of a copular construction.



position. Since it is semantically there, it can be linked to the noun's external theta-role, which is not syntactically present either, in a way proposed by Higginbotham (1985) (see also Wunderlich 1987) for simple adjectives, viz. by theta-identification. So, although this type of modifier has the internal structure of a verbal constituent and denotes an event, it is linked to the noun in an adjectival way and denotes therefore also a property.

This analysis implies that, in our view, there are two ways of modification: direct modification, i.e., linking of theta-roles present in the LCSs of the noun and the modifier, and indirect modification, i.e., modification via a syntactically realized argument (an (empty) operator). Simple adjectives and prenominal (de)verbal modifiers in Dutch and German are related to the noun via direct modification. Relative clauses and postnominal (de)verbal modifiers in English modify the noun indirectly.

But why do (de)verbal modifiers in Dutch and German modify the noun directly but indirectly in English? We claim that argument reduction, and as a consequence direct modification, is the preferred option for economy reasons. In English, however, the Head Final Filter (Williams 1982) excludes participles followed by a complement in prenominal position:

(33) \*the singing a song child

This means that the participial modifier can only be generated as a reduced relative clause in a postnominal position, where it is indirectly linked to the noun via a syntactic argument realized as an empty operator. In Dutch and German, scrambling of the complement circumvents the Head Final Filter and allows for the generation in a prenominal position.

We propose that in English argument reduction is also a possible option. Consider for that matter the following examples with a deverbal, prenominal modifier, which do not violate the Head Final Filter:

(34) the singing child

(35) the (recently) sent book

Both in (34) and (35) the participle has a verbal interpretation: it denotes an event, which in (35) is specified in time. This means that in both cases we are dealing with a verb. Since the participial modifier is in a prenominal position, it is directly linked to the noun via theta-identification, which means that argument reduction has taken place.

In our analysis of (de)verbal modifiers, participles can take an adjectival inflectional ending even if they are verbs, as the following Dutch examples show:

(36) het een liedje zingende kind  
'the child singing a song'

(37) het door Jan gekochte boek  
'the book bought by John'

Although it might be objected that the adjectival inflection shows that the participle is an adjective, there is evidence that modifiers can take an adjectival ending even if they are verbs. In Dutch, infinitives preceded by *te* 'to' can be used as modifiers of the noun:<sup>7</sup>

(38) de door de atleten af te leggen afstand  
'the distance that has to be covered by the athletes'

But whereas Dutch infinitives ending in *-en* never take an adjectival ending (*-e*) in this position (maybe because the infinitival ending is phonetically realized as a schwa, just like the adjectival ending, and can therefore take over its function), infinitives ending in *-n* can do so:

(39) het te ondergane onderzoek  
'the examination (that has) to be undergone'

(40) het te herziene artikel  
'the article (that has) to be revised'

We take this as evidence for our claim that verbs can take an adjectival ending. We assume that the adjectival ending is the head of a functional projection and is associated to the (de)verbal modifier in syntax (see also Vermandere 1998, who adopts such an analysis even for (Germanic) attributive adjectives).

It might also be objected that (de)verbal modifiers cannot be inserted in an adjectival position, unless they have something like a mixed V/A categorial status. We do not think, however, that the pronominal modifier does necessarily have to be adjectival. Nouns combine with predicates, i.e., with [+V] modifiers, which can be adjectives, verbs or (reduced) relative clauses (which

<sup>7</sup> The infinitives in these examples have a deontic reading. *Te* + infinitive can also have an 'ability' reading (*de te voorziene gevolgen* 'the consequences that can be foreseen'). We assume that in that case we are dealing with a *tough*-construction in which e.g., *gemakkelijk* 'easy' has been suppressed.

are, in a sense, extended projections of the verb, see e.g., Grimshaw 1991). These can precede or follow the noun, although the Head Final Filter can block the prenominal position. For English postnominal participial modifiers it is generally assumed that they are verbs (projecting a reduced relative clause structure). We assume the same for prenominal (de)verbal modifiers in Dutch and German, although in our analysis they have undergone argument reduction so that they do not project clausal structure but are directly linked to the noun.

It is true that in the earliest versions of Chomsky's Generative Grammar there were explicit phrase structure rules, which did only mention AP and S' as modifiers of the noun. This might be the reason why linguists have assumed that (de)verbal modifiers have an adjectival or mixed categorial status. But in the Bare Phrase Structure approach of the Minimalist Program (Chomsky 1995), phrase structure is created by means of the operations Merge and Move. In this framework the modifier is adjoined to the noun by means of Merge. The categorial status of the modifier is not important, because it is not inserted in a slot of a special type, e.g., adjectival. It is rather its semantics that determines its possibility to combine with a noun. Only a predicative meaning makes modification of a noun possible.

A final point that we want to discuss in this section is the fact that (de)verbal modifiers are subject to argument reduction. Why would this be so? We propose that the reason for this is that we are dealing with verbal forms that do not express person, number or mood distinctions. They only express tense distinctions. Whereas present participles denote an event in the present, past participles denote an event in the past. Dutch infinitives preceded by *te* in a deontic reading denote a future event:

- (41) *de te schrijven brieven*  
'the letters (that have) to be written'

This inflectional poverty makes the way free for argument reduction. Since person and number are not marked, the subject can be realized as an argument that takes the form of an empty category but can also not be mapped from LCS onto argument structure. The final step results in the absence of the whole syntactic argument structure including the event structure, which gives an adjective (only with participles).

## 5. Deverbal adjectivalization and nominalization

We have argued that deverbal adjectivalization involves a gradual loss of syntactic argument structure and event structure and that only at the last stage can there be a categorial change from a verb into an adjective. Although its categorial status remains verbal during a long time of the process, semantically the verb becomes more and more 'adjectival'. In this section we will discuss three other deverbal processes which involve a gradual semantic change from an event into a property or an object. Whereas in one of these cases, the category also changes only at the last stage of the process as the result of lexicalization, in the other two cases the category already changes at the first stage of the process.

Besides 'adjectivalization' by means of poor verbal inflection ((de)verbal modifiers) there is also 'nominalization' by means of poor verbal inflection, again both participial and infinitival. The English gerund is an example of 'nominalization' by means of participial inflection (42)–(44). The Dutch 'nominalized' infinitive is an example of 'nominalization' by means of infinitival inflection (45)–(47). Both cases involve a gradual process of 'nominalization', which we analyse as the gradual loss of syntactic argument structure and the gradual change of an event into an object:

- (42) John singing this song astonished me.
- (43) John's singing this song astonished me.
- (44) John's singing (of this song) astonished me.
- (45) Fietsen stelen is strafbaar.  
'The stealing of bikes is punishable.'
- (46) Het fietsen stelen (door junkies) neemt steeds meer toe.  
'The stealing of bikes (by junkies) is growing.'
- (47) Het stelen (van fietsen) (door junkies) moet beperkt worden.  
'The stealing (of bikes) (by junkies) must be restricted.'

We assume that at these stages we are still dealing with a verb, and that the process involved is inflection and not derivation, the three stages being fully productive. At the first stage, the verb still has its complete argument structure. One of the arguments can be realized as an empty category (PRO).

At the second stage the inflected form denotes both an event and an object; the external participant role of the verb is only present at the level of LCS and can be realized as an adjunct. At the third stage, the same holds for the internal participant role.

German infinitives can also be 'nominalized', and can even bear nominal inflection:

- (48) die Kunst des Wohnens und Arbeitens in Harmonie mit der Umgebung  
'the art of living-gen. and working-gen. in harmony with the environment'

But in spite of the fact that the 'nominalized' infinitive is written with a capital letter, just like real nouns, and in spite of the nominal inflection, we assume that the 'nominalized' infinitives in (48) are verbs. Just as we did for verbal forms with adjectival inflection, we assume for verbal forms with nominal inflection that the inflection is associated to the verb in syntax and not in the lexicon. If the inflection is the head of a functional projection, probably a Case projection, to which the verbal head is adjoined, verbal forms can take non-verbal inflectional endings without a categorial change being necessary.<sup>8,9</sup>

Verbs with poor inflection can also be lexicalized as nouns, just as participles can be lexicalized as adjectives. Examples of lexicalized participles are result nouns such as *doing(s)* or *ending(s)*. An example of an infinitive that can be lexicalized as a noun is the German infinitive *lächeln* 'to laugh' (*das Lächeln* 'the laugh').

Infinitives and participles can thus both gradually lose their argument and event structure and can both become more and more 'adjectival' or 'nouny'.<sup>10</sup> In their 'adjectival' use, infinitives and participles express, however, different tense distinctions (past and present for participles, future for infinitives). For

<sup>8</sup> Note that whereas in a syntactic affixation approach like Fabb's (1984), a verb can take participial and infinitival inflection in syntax, we assume that verbal inflection, such as participial or infinitival inflection, is a lexical matter. In our approach, only adjectival or nominal (Case) inflection can be added in syntax to poorly inflected verbal forms.

<sup>9</sup> The Case Filter, which in its original formulation (Chomsky 1981, ch. 2) excludes only lexical NPs without Case, would not reject an infinitive without (genitive) Case. However, the Visibility Condition, which replaces the Case Filter (Chomsky 1981, ch. 6), relates Case and  $\theta$ -role assignment. The infinitives in (48) have to be made visible by Case for  $\theta$ -role assignment. Their category is not important.

<sup>10</sup> The last stage is missing in the case of the gradual 'adjectivalization' of infinitives.

'nominalization', languages seem to make a choice between either the participial form (e.g., English gerunds) or the infinitival form (e.g., German and Dutch).

But whereas these are inflectional processes, in our view, there are also derivational deverbal processes. In section 1, we discussed deverbal nominalization, a derivational process that also involves a gradual loss of argument structure. This time the verb changes already at the first stage from a verb into a noun (e.g., *realization*). There is also deverbal adjectivalization via derivation, viz. by means of the derivational suffix *-ble* in English, *-baar* in Dutch and *-bar* in German. Just as in the case of nominalization by means of derivation, we assume that the derivational suffix is category changing. But despite this immediate categorial change we are dealing here too with a gradual loss of syntactic argument structure and event structure and the gradual change of an event into a property. At the first stage, the adjective still has a syntactic argument that can make it function as a reduced relative in English:

(49) rivers [Op navigable by all kinds of ships]

At the second stage, this argument is not mapped from LCS onto syntactic argument structure. The modifier can only be generated in a prenominal position where it is related to the noun via theta-identification, i.e., directly. The following example is from Dutch:

(50) een door muggen overdraagbare ziekte  
'a disease that can be transmitted by mosquitos'

The third stage results in the absence of syntactic argument structure and event structure (see also Fabb 1984 and van der Putten 1997 for the distinction between verbal forms, which are syntactically derived in their theory, and lexicalized forms):

(51) an adorable child

We conclude that both inflectional and derivational (de)verbal processes involve a gradual change in syntactic argument and event structure. This is illustrated by the following table:

## (52) (De)verbal processes

|                   | inflection                 | derivation                 |
|-------------------|----------------------------|----------------------------|
| adjectivalization | participles<br>infinitives | adjectives (-ble, -ba(a)r) |
| nominalization    | participles<br>infinitives | action nouns (e.g., -ion)  |

We have argued that both inflectional 'adjectivalization' and 'nominalization' and derivational adjectivalization and nominalization are gradual processes in which there is a gradual loss of syntactic arguments and finally of event structure. The main difference between the inflectional and the derivational processes lies in the first step of the gradual change. In the case of the inflectional processes there is first no change in syntactic argument structure and category (in *the children singing the song*, the present participle has exactly the same argument structure and category as the active verb *to sing*), whereas in the case of the derivational processes there is immediately a change in argument structure (externalization of the internal argument in the case of adjectivalization and internalization of the external argument in the case of nominalization) and category.

Haspelmath (1996) assumes that both 'verbal' prenominal participial modifiers in German and gerunds in English involve a categorial change. On the basis of this he argues that (participial) inflection can be category changing, contrary to what is generally claimed in the literature. In this paper we defend the idea that the first stages of gradual inflectional (de)verbal processes do not involve a categorial change. It is only at the last stage that the category can change from a verb into an adjective or a noun. This last process is not productive and the resulting form can have an idiosyncratic meaning, properties normally ascribed to lexicalization. This means that the idea that inflection does not change the category can be maintained.

## 6. Conclusion

In this paper we have argued that prenominal present and past participles are either lexicalized adjectives or event-denoting verbs with a reduced syntactic argument structure. The eventive modifiers have the internal structure of verbal constituents, but behave externally like attributive adjectives, which do not project a syntactic argument either. Both prenominal participles denoting

an event and adjectives are linked to the noun in a direct way, via theta-identification. Adjectival inflection is associated to the verbal form in syntax.

Unlike English, Dutch and German do not have deverbal modifiers with complete argument structure, that is, clausal participial modifiers, because they have a more economical way of modification at their disposal, thanks to the availability of scrambling. We have argued that deverbal adjectivalization is a gradual process that derives from a verbal form with its complete syntactic argument structure an intermediate form with partially verbal and partially adjectival properties and finally a purely adjectival form.

We have compared inflectional and derivational deverbal adjectivalization and nominalization and have claimed that only the derivational processes involve an immediate categorial change, followed by a gradual loss of argument structure and event structure. This gradual loss also takes place in the case of the inflectional processes, but this time the categorial change is the outcome of lexicalization as the last stage of the process.

### References

- Bennis, H. – Wehrmann, P. 1990. On the categorial status of present participles. In: Bok-Bennema, R. – Coopmans, P. (eds): *Linguistics in the Netherlands 7*: 1–11.
- Borer, H. 1990. *V + ing*: It walks like an adjective, it talks like an adjective. In: *Linguistic Inquiry* 21: 95–103.
- Brekke, M. 1988. The experiencer constraint. In: *Linguistic Inquiry* 19: 169–80.
- Chomsky, N. 1981. *Lectures on government and binding*. Foris, Dordrecht.
- Chomsky, N. 1995. *The minimalist program*. The MIT Press, Cambridge MA.
- Fabb, N. 1984. *Syntactic affixation*. Ph.D. dissertation. MIT, Cambridge MA.
- Fanselow, G. 1986. On the sentential nature of prenominal adjectives in German. In: *Folia Linguistica* 20: 341–79.
- Grimshaw, J. 1990. *Argument structure*. The MIT Press, Cambridge MA.
- Grimshaw, J. 1991. *Extended projection*. Unpublished paper. Brandeis University.
- Haspelmath, M. 1996. Word-class changing inflection and morphological theory. In: Booij, G. – Marle, J. van (eds): *Yearbook of Morphology 1995*, 43–66. Kluwer, Dordrecht.
- Higginbotham, J. 1985. On semantics. In: *Linguistic Inquiry* 16: 547–93.
- Hoekstra, T. 1984. *Transitivity: grammatical relations in Government-Binding theory*. Foris, Dordrecht.
- Hout, A. van 1991. Deverbal nominalization, object versus event denoting nominals: implications for argument and event structure. In: Drijkoningen, F. – Kemenade, A. van (eds): *Linguistics in the Netherlands 8*: 71–80.



- Levin, B. – Rappaport, M. 1986. The formation of adjectival passives. In: *Linguistic Inquiry* 17: 623–61.
- Putten, F. van der 1997. Matter and Mind in morphology. Syntactic and lexical deverbal morphology in Dutch. Ph.D. dissertation. Universiteit van Leiden.
- Riemsdijk, H. van 1981. The case of German adjectives. In: *UMass Occasional Papers in Linguistics* 6: 148–73.
- Sadler, L. – Arnold, D. 1994. Prenominal adjectives and the phrasal/lexical distinction. In: *Linguistics* 30: 187–226.
- Siegel, D. 1973. Non-sources for unpassives. In: Kimball, J. (ed.): *Syntax and Semantics II*, 301–17. Seminar Press, New York.
- Sleeman, P. – Verheugd, E. 1998. How reduced are reduced relatives? In: Kager, R. – Bezooijen, R. van (eds): *Linguistics in the Netherlands* 15: 187–99.
- Toman, J. 1986. A (word-)syntax for participles. In: *Linguistische Berichte* 105: 367–408.
- Vermandere, D. 1998. Projecting DP: a view from the (Dutch) adjective. Preprint. University of Leuven.
- Williams, E. 1981. Argument structure and morphology. In: *The Linguistic Review* 1: 81–114.
- Williams, E. 1982. Another argument that passive is transformational. In: *Linguistic Inquiry* 13: 160–3.
- Wunderlich, D. 1987. Partizipien in Deutsch. In: *Linguistische Berichte* 111: 345–417.

Address of the authors: Petra Sleeman and Els Verheugd

French Department  
University of Amsterdam  
Spuistraat 134  
1012 VB Amsterdam  
The Netherlands  
p.sleeman@hum.uva.nl  
els.verheugd@hum.uva.nl



## INFLECTION AND THE LEXEME\*

ANDREW SPENCER

### Abstract

The notion of 'lexeme' is central to realizational theories of morphology and to the notion of 'morphology by itself'. It is generally assumed that 'inherent' inflections such as Plural or Past Tense impart a meaning to the inflected word. However, this runs counter to the usual understanding of the notion 'lexeme', which is supposed to have a single constant meaning for all word forms. Since derivational morphology is supposed to create new lexemes by adding a new component of meaning this makes it difficult to distinguish inflection from derivation, which in turn undermines the whole lexeme concept. The problem evaporates if we assume that syntactic features are properties of phrases, not individual words, and that their semantic interpretation is therefore defined over phrasal expressions. This brings inflections on lexical heads into line with all other exponents of functional features.

### 1. The lexeme concept

All theories of morphosyntax appeal to a notion of 'lexical entry', a representation of formal and conceptual properties of a word, distinguishing it from other words and capturing those unpredictable properties which govern its morphological and syntactic behaviour. For morphologists who adopt a 'realization' approach to inflection (and possibly derivation), the lexical entry is the

\* I gratefully acknowledge the support of the Economic and Social Research Council (Project R000236115) in the preparation of sections of this paper. In addition to being presented at the 8th International Morphology Meeting, some of this material has been presented at the Linguistics Association of Great Britain, Autumn Meeting, University of Hertfordshire, 4-6 September 1997, the Department of English and American Studies, University of Coimbra, 3 October 1997, the seminar *Challenges for Inflectional Description-4*, School of Oriental and African Studies, 31 October 1997 (funded by the Economic and Social Research Council), and at Oxford University, 14 November 1997. I am grateful to the audiences for a number of helpful suggestions and criticisms.

'lexeme'. A standard characterization of 'lexeme' following the tradition inaugurated by Matthews (1972) and developed by Aronoff (1994), Beard (1995) and others would run something like (1):

- (1) Lexeme: Set of words forms united by constant meaning

This characterization captures two important claims: (i) a lexeme has a single meaning, (ii) the lexeme generalizes over inflected forms. This notion of lexeme hinges on the distinction between inflected forms of words (grammatical words), as opposed to derived lexemes. That is, the lexeme concept presupposes the distinction between inflection and derivation. A derivational process will create a new lexeme, which minimally ought to mean that it adds some element of meaning to the base lexeme.<sup>1</sup> If the basic difference between lexemes lies in conceptual meaning then it will be important to distinguish between the meanings conveyed by inflections and those conveyed by derivation. For instance, we do not want to say that the form *dogs* realizes a different lexeme from the form *dog* on the grounds that *dog* means [DOG] while *dogs* means [PLURAL[DOG]]. Yet it is a common assumption that inflections add meaning to words. Anderson (1992, 79) explicitly states that *dog* and *dogs* have different meanings, and Booij (1994; 1996) draws a distinction between inherent inflections, those inflections which bear meanings in themselves such as Number or Tense, and contextual inflections, those inflections which are purely formal and are determined purely by the syntactic context (agreement, government). It is clear from Booij's characterization that inherent inflections are properties of the word.

If inflected words may bear the meaning of those inflections, then just looking at the semantics of a word will not tell us whether a morphological process that affects it is inflectional or derivational. Clearly we need a way of distinguishing the properties of inflections from those of derivations. A checklist of such properties is provided by Plank (1994). However, he argues that the distinction is, in fact, gradient (cf. Bybee 1985; Haspelmath 1996). For Anderson (1992) inflection and derivation are handled in different components

<sup>1</sup> In fact, the discussion of derivation here is a massive oversimplification, since there is derivation which is entirely divorced from semantic interpretation as well as morphology which appears to be derivational but which adds no new meaning, i.e., transpositions such as participles, action nominalizations and gerunds ('masdars') (see Haspelmath 1996) and also relational adjectives (*morphology* ~ *morphological*), and perhaps property nominalizations of adjectives (*sweet* ~ *sweetness*). See Spencer (1999) for proposals for treating transpositions.

of grammar (syntax and the lexicon respectively), and we will see that this is an important insight. However, this solution still leaves open the problem of how we know (or rather, how the language learner knows) that [PLURAL[DOG]] represents a word form of a single lexeme DOG and not a distinct lexeme (created by the syntax).

It is important to understand that inflectional morphology is just one way in which a language realizes its f(unctional) features. I shall assume that every language has a vocabulary of f-features, which have two main roles: (i) to regulate syntactic structure (e.g. agreement, government) and (ii) to express certain 'functional' meanings, such as Tense, Definiteness, Plural, etc. In the context of inflection these two roles correspond respectively to contextual inflection and inherent inflection in the sense of Booij (1994; 1996). F-features may be expressed in a great variety of ways apart from inflection of the head word, including clitics, auxiliary words, word order and intonation, or by a combination of these. Intonation extends over phrases and presumably can only be interpreted on the basis of a parsing of the intonation phrase. Word order by definition is defined over sequences of words. In many syntactic frameworks auxiliary elements such as auxiliary verbs, determiners, degree modifiers and so on are regarded as heads of functional phrases which take lexically headed phrases as complements. In those theories, the semantic interpretation of auxiliary words must therefore by definition include the lexical phrase in its scope. Clitics and edge inflections are generally defined with respect to a position at the periphery of a phrase. For instance, the possessive *-s* in English appears at the right edge of a determiner phrase, even though it relates semantically to the lexical head of the noun phrase (e.g., [[*the man we were talking to*]'s] *name*). Such cases of phrasal affixation are generally analysed as peripheral marking serving as the exponent of a phrase level feature (e.g., Halpern 1995). This means that if head-inflections such as Plural or Past are interpreted at the word level rather than the phrasal level they are unique amongst f-features.

## 2. Inflection and the architecture of the lexicon

In any theory it is necessary to establish a correspondence between syntactic f-features and (i) the realization of those features; (ii) the semantic interpretation of those features. In many cases there is a reasonably clear correlation between morphology and syntactic f-features. Thus, if we regard English modals as exponents of mood f-features we can say that there is a feature,



In (2) we see that the syntactic category label is given an identical right and a left subscript cross-referencing a semantic and a morphological representation respectively. However, this only accounts for an underived, uninflected lexical entry. Now consider the inflected forms, specifically the plural. In a realization theory the plural would be specified by a function or set of functions applying to the lexeme (cf. Aronoff 1994) or to the root of the lexeme (Stump's (1993) Paradigm Functions). Call the function which delivers plural forms PLUR. Suppose that this function is coindexed with a syntactic feature [NUM:Pl]. Using Jackendoff's notation we could represent this as in (3):

$$(3) \quad \begin{array}{ccc} {}_{10}\text{PLUR}({}_1\text{dog}) & {}_{10}[\text{NUM:Pl}]_{10}({}_1[\text{N}]_1) & [\text{Thing}^{\text{PLUR}}[\text{Thing}^{\text{DOG}}]_1]_{10} \\ | & & \\ \text{dog-}z & & \end{array}$$

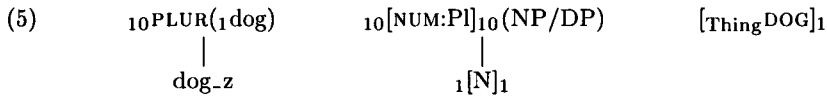
I have taken the liberty of interpreting the functional feature here as a function applied to the lexical category feature in the syntax, to mirror the functional application in the morphological and semantic representations.

But as we know, (3) is odd because it seems to be saying that inflection creates a new lexeme. Given the logic of the lexeme concept we should really be assuming the set of representations in (4):

$$(4) \quad \begin{array}{ccc} {}_{10}\text{PLUR}({}_1\text{dog}) & {}_{10}[\text{NUM:Pl}]_{10}({}_1[\text{N}]_1) & [\text{Thing}^{\text{DOG}}]_1 \\ | & & \\ \text{dog-}z & & \end{array}$$

In other words, *dogs* is a form of the DOG lexeme which bears an exponent which is (somehow, somewhere) interpreted as 'plural'. Note that we need representations of this sort anyway for Booij's contextual inflections. Given (4), *dogs* differs minimally from *the dog*, which is also a morphosyntactic expression that contains an exponent of the DOG lexeme (note, not the THE DOG lexeme!) in addition to an exponent of the feature 'definite'. The difference between the plural and the definite form is simply that the plural is expressed by means of a Paradigm Function applied to the root of the lexeme, while the definite form is expressed by means of an auxiliary element (definite determiner). In languages with definiteness affixes rather than definite articles the parallel between a definite dog and a plural dog would be complete. Semantically speaking, plurality cannot be a property of a bare common noun but only of a referential element. The syntactic unit corresponding to a referential element is a noun (or determiner) phrase. The obvious way to capture this is to say that the syntactic plural feature is a property of the noun phrase, or

perhaps even determiner phrase, where it receives its semantic interpretation (*pace* Anderson 1992). This view of the word form is shown in (5):



The meaning of the plural marker (or indeed, whether it gets interpreted semantically in the first place) depends on the wider syntactic context. The correspondence rules ensure that the PLUR function in the morphology licenses the appearance of the [NUM:PI] feature on the noun/determiner phrase in the syntax. If nothing else is said, the coindexation will ensure that the word form which realizes the [NUM:PI] feature is the lexical head (other machinery is necessary to account for edge inflections, as argued by Halpern 1995).

There are several reasons for marking and interpreting inflections at the phrasal level only.

First, it brings head-inflections into line with all other exponents of f-features (including edge inflections).

Second, there are occasions when it is quite unclear what the additional meaning component of *dogs* actually is, if it exists at all, as in the expression *two dogs*. The simplest representation for this will be something along the lines of  $[2x[\text{DOG}(x)]]$ , without any semantic interpretation given to the syntactic plural feature, which is therefore interpreted as a kind of agreement. Not all languages insist on this type of agreement, of course (e.g., Hungarian), and in English agreement is not found with nouns denoting game animals (*We bagged a brace of pheasant/\*pheasants*). In point of fact, it might be embarrassing if the plural marker were given semantic interpretation in *two dogs*, because then the expression ought to mean 'two groups of dogs'. Now, there are no doubt ways of defining the semantics of plurality in such a way as to avoid this (see Ortmann 1998, for instance), but if the syntactic feature is phrasal and if the syntax-semantic correspondence principles recognize it as an agreement, the simplest interpretation of plurality can be maintained.

There are other constructions in which the plural fails to receive semantic interpretation. Thus, in (6a, b) we have two different ways of saying essentially the same thing:



- (6) (a) these types of dogs  
 (b) these types of dog  
 (c) this type of dog  
 (d) \*this type of dogs

In (6a, c) it appears as if there is some sort of number agreement between *type* and *dog*, which is not, however, obligatory ((6b) with a plural head noun). Likewise, in predicative constructions such as (7) we find number agreement:

- (7) (a) None of her daughters were doctors  
 (b) None of her daughters was a doctor  
 (c) \*None of her daughters were a doctor  
 (d) \*None of her daughters was doctors

In (7a, b) we see that *none* can be treated as singular or plural, but once that decision is made it must be carried through for the whole predicate (cf. (7c, d)). However, since (7a, b) are synonymous it can hardly be the case that the plural morphology contributes meaning (to the lexeme or to anything else).

The past tense inflection of English verbs is if anything even more problematical, for well-known reasons. Thus, given sequence of tenses, we don't want to have to say that *arrived* in (8) means [PAST[ARRIVE(*x*)]]:

- (8) I thought you arrived tomorrow.

Here, the past tense marking is conditioned by agreement, and thus constitutes a semantically uninterpreted morphosyntactic feature. The last thing we want *arrived* to mean in (8) is [PAST[ARRIVE(*x*)]].

Constructions such as the English perfect tense pose interesting problems when we come to ask about semantic interpretation. In an expression like *Tom has eaten the apple* the meaning of 'perfect' is expressed neither by *has* nor by *eaten*. This is because the participle is generally ambiguous between the perfect and the passive reading (at least for transitive verbs) and *have* is ambiguous between a main verb, a modal auxiliary (*Tom has to leave*), and, arguably, a meaningless 'light' verb (*Tom had a bath*). What conveys the meaning 'perfect' is the combination of auxiliary and participle. This is different from a case such as *Tom must leave* in which we can identify the modal force of the sentence simply with the auxiliary. We can think of the modal auxiliary as an example of a compositional analytical construction, and the perfect tense as a non-compositional analytical construction. An interesting question (examined at

some length in Ackerman–Webelhuth 1998) is the relationship between such analytical constructions on the one hand and synthetic constructions like the simple past (*Tom ate an apple*) on the other. In one sense, these are in a paradigmatic relation to each other, and the perfect or progressive forms are part of the inflectional paradigm of the verbal lexeme. This is particularly obvious in cases of periphrastic inflection of the kind discussed for Latin by Börjars–Vincent–Chapman (1997).

The reason this is of interest is that it is very difficult to see how to state the paradigmatic relationship between synthetic constructions and analytic constructions given the standard conception of inflection. This is because classical inflection is only a part of the story. A full grammar of English has to set correspondence rules between syntactic features such as [ASPECT:Perfect, TENSE:Present] and 'have + past participle' constructions. The correspondence rules which tell us that the morphological function PAST(*leave*) (i.e., *left*) corresponds to the syntactic feature [TENSE:Past] applied to LEAVE will also tell us that *has left* is the construction required to realize the syntactic feature specification [ASPECT:Perfect, TENSE:Present].

There remains one question: how does the grammar differentiate between uninterpreted inflections and interpreted inflections? Note that this is a problem for any theory of grammar, though not one which has received much discussion. (Booij (1994; 1996) does not explicitly address the problem, for instance, in his discussion of contextual and inherent inflection.) The solution depends on the overall architecture of the grammar. One suggestion we might make is the following. Suppose we follow Lexical Functional Grammar (Bresnan, in press) in assuming a level of c-structure defining the phrasal constituency of an expression and a level of f-structure at which functional properties such as grammatical relations, tense and so on are expressed. In standard LFG, this level also contains feature specifications governing agreement and government (contextual inflection). However, suppose we say that all inherent f-features are represented twice, once in f-structure and also as node labels at c-structure. Suppose, too, that contextual features are represented only at c-structure, being distributed there by specific rules of formal agreement and government. Finally, suppose that all semantic interpretation is defined over f-structure (as in standard LFG). In this fashion we can make a simple architectural distinction between the two types of feature. This is not the standard practice in LFG, though it is not incompatible with the basic precepts of the theory as far as I can tell.

### 3. Conclusions

The model argued for here shares much with Anderson's (1992) conception of 'split morphology'. However, Anderson argues for an architecture in which inflections are actually specified in the syntax. In my model, derived from the perspective on the lexicon offered in Jackendoff 1997, it is possible to accommodate all morphology, derivational or inflectional, in an autonomous morphological module, while still enjoying the benefits of split morphology. I adopt Jackendoff's suggestion that the components of a word are linked to each other by means of a set of lexical indices. In simple cases there is a one-one correspondence between features at different levels, though in more complex cases we might find a more complex relationship (as in the case of the English perfect or Colloquial French *passé composé*). The split between inflection and derivation is apparent from the correspondence patterns of features at various levels of representation. A piece of morphology will be interpreted as inflectional if and only if it corresponds to a syntactic feature. Syntactic features, if interpreted at all, are interpreted at the phrasal level. This helps explain why inherent inflections such as the English plural sometimes behave as though they were contextual inflections, in not receiving a semantic interpretation. In this way inflections are seen to be no different from other functional features (though we are not obliged to say that inflections are syntactic heads projecting their own phrases, as in most Principles and Parameters models of syntax). On the other hand a piece of morphology is derivational if it is given direct semantic interpretation at the lexemic level.

As pointed out by opponents of split morphology, there is no formal difference between morphological operations subserving inflection and those subserving derivation (indeed, one and the same operation can subserve both at the same time for the same lexeme, as we find with, say verbal and adjectival passives in English). However, this is expected in a theory which countenances an autonomous morphological module and the principle of Separationism, as here.

### References

- Ackerman, F. – Webelhuth, G. 1998. A theory of predicates. Center for the Study of Language and Information, Stanford University.
- Anderson, S.R. 1992. A-morphous morphology. Cambridge University Press, Cambridge.
- Aronoff, M. 1994. Morphology by itself. MIT Press, Cambridge MA.

- Beard, R. 1995. *Lexeme morpheme base morphology*. SUNY Press, Stony Brook NY.
- Booij, G. 1994. Against split morphology. In: Booij, G. – van Marle, J. (eds): *Yearbook of Morphology 1993*, 27–49. Kluwer, Dordrecht.
- Booij, G. 1996. Inherent versus contextual inflection and the split morphology hypothesis. In: Booij, G. – van Marle, J. (eds): *Yearbook of Morphology 1995*, 1–16. Kluwer, Dordrecht.
- Börjars, K. – Vincent, N. – Chapman, C. 1997. Paradigms, periphrases and pronominal inflection: a feature-based account. In: Booij, G. – van Marle, J. (eds): *Yearbook of Morphology 1996*, 155–80. Kluwer, Dordrecht.
- Bresnan, J. in press. *Lexical Functional Syntax*. Blackwell, Oxford.
- Bybee, J.L. 1985. *Morphology: A study of the relation between meaning and form*. John Benjamins, Amsterdam.
- Halpern, A. 1995. On the placement and morphology of clitics. CSLI Publications, Stanford.
- Haspelmath, M. 1996. Word-class-changing inflection and morphological theory. In: Booij, G. – van Marle, J. (eds): *Yearbook of Morphology 1995*, 43–66. Kluwer, Dordrecht.
- Jackendoff, R.S. 1997. *The architecture of the language faculty*. MIT Press, Cambridge MA.
- Matthews, P. 1972. *Inflectional morphology*. Cambridge University Press, Cambridge.
- Ortmann, A. 1998. Where plural refuses to agree: feature unification or not? Talk given at the 8th International Morphology Meeting, Budapest, 11–14 June 1998. [Also in this volume]
- Plank, F. 1994. Inflection and derivation. In: Asher, R.E. (ed.): *The Encyclopaedia of Language and Linguistics*, Vol. 3, 1671–8. Pergamon Press, Oxford.
- Spencer, A. 1997. A note on the lexeme and the paradigm. *Essex Research Reports in Linguistics* 16: 1–16.
- Spencer, A. 1999. Transpositions and argument structure. In: Booij, G. – van Marle, J. (eds): *Yearbook of Morphology 1998*, 72–102. Kluwer, Dordrecht.
- Stump, G.T. 1993. Position classes and morphological theory. In: Booij, G. – van Marle, J. (eds): *Yearbook of Morphology 1992*, 129–79. Kluwer, Dordrecht.

Address of the author: Andrew Spencer  
 Department of Language and Linguistics  
 University of Essex  
 Colchester CO4 3SQ  
 United Kingdom  
 spena@essex.ac.uk

## MORPHOLOGICAL PROPERTIES IN THE LEXICON: DIACHRONIC EVIDENCE

WOLFGANG ULLRICH WURZEL

### Abstract

The paper tries to demonstrate how it is possible to gain evidence for the detailed structure of grammatical systems by investigating language change. The decisive methodological question is: How must a system be structured so that a certain change can take place? The paper presents several morphological changes from the history of Germanic languages, on the basis of which some important questions concerning the structure of morphological lexicon representations can be answered in a justified manner. Altogether, it is intended to contribute to the defeat of the fruitless strict separation of synchrony from diachrony in linguistics.

### Introduction

The results presented in this paper are part of a project whose goal is to draw conclusions from the processes of language change about the grammatical structure that is present at the onset of the particular changes. Thus, structure is inferred from subsequent change; synchrony is inferred from diachrony, thereby reversing the more traditional method of inferencing change from a preexisting structure, which is characterized by the question of what changes are possible in a given system. In this sense, it follows the line of Kiparsky (1968, 174), where language change is characterized as “a window on the form of linguistic competence that is not obscured by factors like performance”. The fundamental question in our approach is, what are the structural prerequisites for a specific change that has been documented historically? Based on the study of language changes seen from this angle and on corresponding generalizations, requirements will be developed which grammatical theories must satisfy if their claim to explanation is to include historical variability as one of the fundamental, universal characteristics of natural language, i.e., if they are to be adequate in a diachronic sense.

In our approach, we make two important assumptions: First, only changes that are grammatically determined, i.e., are “set by the system” will be relevant

in this context, and not socially motivated changes, and I mean socially in the broadest sense. Second, we make the plausible, empirically upheld assumption that grammatically motivated change always leads to the decreasing of markedness, i.e., to the formation of locally preferred grammatical structures (which, at the same time, explains why such changes occur in the first place).<sup>1</sup>

In the remainder of the paper, I will discuss the following four aspects of morphological lexicon representation:

1. The formal representation of lexemes in the lexicon
2. Types of inflectional class membership and their representations
3. Factors in determining inflectional class membership
4. The specification of marked inflectional class membership.

Although I am only able to discuss a small number of examples in this short paper, for each example I cite, there are enough parallel cases from various languages to insure that the hypotheses formulated are well supported by factual evidence.

### 1. The formal representation of lexemes in the lexicon

There are two opposing views on the representation of lexemes in the lexicon, both of which are formulated within a framework that does not even take the lexical storing of all forms of a lexeme into account, and both of which are supported in turn by variously motivated synchronic arguments. The first is the representation of lexemes in the form of a morpheme, specifically the base morpheme, and the second is the representation in the form of one or more concrete word-forms. The morphology is conceptualized accordingly as either morpheme-based or word-based in this sense. Let us examine a particular

<sup>1</sup> This assumption is not uncontroversial. The main reason for this seems to be that markedness is often understood as “markedness per se”. But markedness only exists as “markedness regarding a given parameter”. This means that a linguistic unit is more or less marked only relative to a given parameter. Therefore, “markedness reduction” always means “reduction of markedness regarding a given parameter” and the decrease of markedness by change regarding parameter X may condition the increase of markedness regarding parameter Y (cf. Vennemann 1989; Wurzel 1994, 28ff). Thus, for instance in inflectional morphology, a change that reduces markedness regarding inflection class membership (the only type of markedness treated in this paper, cf. section 2) may easily lead to more markedness regarding constructional iconicity. Cf. the class transitions of nouns in *-er* like *junker* ‘squire’ and *adler* ‘eagle’ from the (weak) *n-* to the (strong) *a-*declension in Early New High German: *junker* – pl. *junker-n* > *junker* – pl. *junker*, where the category of plural is no longer symbolized by a marker on the word after the change.

language development that sheds some light on which of these two assumptions is correct, that is, on how speakers actually store lexical units.

**Example 1: Transfer of masculine nouns from the weak to the strong declension in German**

Old High German weak masculine nouns always end with the vowel *-o*, e.g., *boto* ‘messenger’. The reflex of this *-o* in Middle High German is an *-e*, e.g., *bote*. Yet some of these words lose their final vowel already in Middle High German due to phonological deletion, and this development continues in later times. The words affected by the deletion therefore end in a consonant, a diphthong or a long vowel (phonologically in a branching rhyme); cf. *helm* ‘helmet’. As such, they have the same phonological word endings as the the strong *a*-declension masculines; cf. *halm* ‘stalk’. Beginning with the loss of final vowels, these nouns tend to shift from the weak masculine class with *n*-plural formation to the strong masculine class with *e*-plural formation, cf. for example Middle High German class transfers such as *helm*, *hirz* ‘stag, male deer’, *kern* ‘seed, kernel’ – gen.sg. *des helmen*, *hirzen*, *kernen* – nom.pl. *die helmen*, *hirzen*, *kernen* > *helm*, *hirz*, *kern* – gen.sg. *des helmes*, *hirzes*, *kerns* – nom.pl. *die helme*, *hirze*, *kerne*. Later transfers include among others *blitz* ‘lightening’, *greis* ‘old man’, *pfau* ‘peacock’, *schwan* ‘swan’ – gen.sg. *des blitzen*, *greisen*, *pfauen*, *schwanen* – nom.pl. *die blitzen*, *greisen*, *pfauen*, *schwanen* > *Blitz*, *Greis*, *Pfau*, *Schwan* – gen.sg. *des Blitzes*, *Greises*, *Pfaus*, *Schwans* – nom.pl. *die Blitze*, *Greise*, *Pfaue*, *Schwäne*.

Let us look at the two possibilities of lexical storage. If lexical storage is morpheme-based, we have representations as in (1), if it is word-based, representations are as in (2); stage I is before and stage II is after the *e*-deletion (BM meaning base morpheme, BM/N nominal base morpheme):

- |     |           |     |                                       |     |                                      |
|-----|-----------|-----|---------------------------------------|-----|--------------------------------------|
| (1) | stage I:  | (a) | [helm] <sub>BM/N</sub>                | (b) | [halm] <sub>BM/N</sub>               |
|     | stage II: | (a) | [helm] <sub>BM/N</sub>                | (b) | [halm] <sub>BM/N</sub>               |
| (2) | stage I:  | (a) | [[helm] <sub>BM</sub> e] <sub>N</sub> | (b) | [[halm] <sub>BM</sub> ] <sub>N</sub> |
|     | stage II: | (a) | [[helm] <sub>BM</sub> ] <sub>N</sub>  | (b) | [[halm] <sub>BM</sub> ] <sub>N</sub> |

If we assume that the words are represented in the form of morphemes as in (1), it cannot be explained that the class transfers begin when the word final vowel *-e* has disappeared, because this vowel is not part of the lexical representation. The conditions for the morphological class change would be equally good before and after the vowel deletion. The relation between vowel deletion and class change would merely be accidental. The morphological change will only be

explainable if the lexemes are represented as in (2), in the form of nom.sg., the base form of the paradigm. The onset of morphological change is dependent on the phonological form of the word, specifically on the form of the nom.sg. The change begins at the point when speakers can no longer identify the lexeme as a member of the former class. As long as masculine nouns still end in *-e*, they cannot belong to the *a*-class. Because of this, no class shifting occurs before phonological neutralization. The transfers begin only after the form of the nom.sg. has become ambiguous with respect to its inflectional class, that is, when the weak masculine *helm* and the strong masculine *halm* are no longer distinguishable in this regard. At this point the larger class in each case has the stronger “attracting power” (we will return to this point later). In other words, as long as there is an overt nom.sg. morpheme, it functions as an indicator of inflectional class.

These facts invite the conclusion that lexemes in the lexicon are not represented in the form of base morphemes, but as concrete word forms. As shown in our example, phonological changes of the nom.sg. lead to changes in the inflected forms of the words (even if this does not happen in all cases, there is a clear tendency). Based on this observation, we may conclude further that not all inflected forms of paradigms are stored lexically either; rather, for each lexeme, only specific members of the paradigm are stored, which we call base forms. The base form of the German noun is, as the example shows (and is to be expected), the nom.sg.<sup>2</sup>

Thus, our first hypothesis concerning lexical representation is as follows:

- (A) Lexemes are represented in the lexicon as concrete base forms. As such, both the representation of all inflected forms and the representation in the form of base morphemes are excluded. In this sense, inflectional morphology is word-based and not morpheme-based.<sup>3</sup>

## 2. Types of inflectional class membership and their lexical representations

Belonging to a particular inflectional class is a morphological property of the individual lexeme. This property must therefore be specified in inflectional systems with competing inflection classes in the lexical representations of the

<sup>2</sup> Less trivially, the base form of the verb, at least in German, is the infinitive and not the 3sg.pres.ind., as relevant changes show; cf. Wurzel (in prep.).

<sup>3</sup> Of course the base form may consist of a single morpheme, cf. MHG *halm* and most NHG nouns like *Hund* or *Frau* (vs. *Bot-e* and *Katz-e*).



lexemes. The question as to how this happens will be discussed in this section with reference to relevant language changes.

In this new context, let us return to example 1 discussed above, the transfer of masculine nouns from the weak to the strong class in German. This change is characterised by the fact that the prerequisites for a morphological change, i.e., a change in inflectional class, are created by phonological changes that affect the phonological structure of the base form; cf. again MHG stage I  $[[\text{helm}]_{\text{BM}} \text{e}]_{\text{N}} > \text{MHG stage II } [[\text{helm}]_{\text{BM}}]_{\text{N}}$ , whereby the affected weak masculines are no longer formally distinguishable from the strong masculines, for example the strong masculine *halm* with the representation  $[[\text{halm}]_{\text{BM}}]_{\text{N}}$ . But why do transfers in inflectional class really occur in such cases?

We assumed above that grammatically initiated change always decreases grammatical markedness, which indeed motivates such change in the first place. This invites the question as to the nature of the decrease in markedness in inflectional class shifts as in our example. Even at stage II, where speakers can no longer determine from the respective base forms how the lexemes are to be inflected, nouns such as *helm* and *halm* must be represented differently in the lexicon if they are to be assigned to different inflectional classes. What do the representations at this stage look like? There are two possibilities basically: either both types of words, the weak and the strong masculines, are specified as members of inflectional classes, or only one type of words is inflectionally specified, specifically those which belong to what is for them the less normal inflectional class. That would clearly be the weak nouns in this case, because the strong masculine class contains considerably more words that end in consonants, diphthongs or long vowels than the weak class. Thus in the first scenario we would have the lexical representations in (3) and in the second scenario the lexical representations in (4), where CS stands for class specification:<sup>4</sup>

(3) (a)  $[[\text{halm}]_{\text{BM}}]_{\text{N}}: [\text{CS Y}]$                       (b)  $[[\text{helm}]_{\text{BM}}]_{\text{N}}: [\text{CS X}]$

(4) (a)  $[[\text{halm}]_{\text{BM}}]_{\text{N}} \emptyset$                                       (b)  $[[\text{helm}]_{\text{BM}}]_{\text{N}}: [\text{CS X}]$

It is easy to see that the transfer of words like *helm* from the weak to the strong class is only compatible with the representation in (4), and not with the representation in (3). A transfer  $[[\text{helm}]_{\text{BM}}]_{\text{N}}: [\text{CS X}] > [[\text{helm}]_{\text{BM}}]_{\text{N}}: [\text{CS Y}]$  shows no decrease in markedness, that is, in grammatical complexity, while a

<sup>4</sup> These are intended as makeshift inflectional class specifications; cf. the discussion under examples 4 and 5.

change  $[[\text{helm}]_{\text{BM}}]_{\text{N}}: [\text{CS X}] > [[\text{helm}]_{\text{BM}}]_{\text{N}}$  does, in that the inflectional class specification is deleted.

In other words, language changes like the one just discussed clearly show that, in cases where speakers cannot tell for certain how a word will inflect, where in this sense we have competing inflectional classes, only the less normal inflectional class memberships are specified in the lexicon.<sup>5</sup> Thus, the specification of inflectional classes follows the default principle. Otherwise, such changes would not be motivated and their occurrence would not be explainable. The affected words always transfer from the classes in which they are marked, lexically specified members to classes in which they are unmarked, lexically unspecified members.

From this we formulate the following two hypotheses:

- (B) If the inflectional class membership of a lexeme is not obvious due to its independent properties, then the lexeme usually has an unmarked class membership. Consequently, even in such cases, inflectional class membership is not simply arbitrary.
- (C) In the lexicon, only lexemes with marked class membership are specified for their inflectional properties; lexemes with unmarked class membership remain unspecified. Thus, the specification of inflectional class is minimalistic; it follows the default principle.<sup>6</sup>

### 3. Factors in determining inflectional class membership

The example discussed shows that inflectional class can depend on phonological properties of the base forms. But phonological properties are not the only determining factor of inflectional class. The next example shows that syntactic properties of the lexeme may also be relevant in this regard:

#### **Example 2: Noun class transfers triggered by gender changes in German**

German feminine nouns ending in *-el* (like those ending in *-er*) usually inflect according to the weak paradigm, e.g., *die Amsel* 'blackbird', *Insel* 'island',

<sup>5</sup> One could imagine a third possibility also fitting the assumption that grammatically conditioned change goes from more marked to less marked: Both types of inflectional class membership are stored in the lexicon explicitly but only the marked one "strains" the lexicon. As I think, this would be a notational variant of the second possibility in (4) having the two disadvantages that morphological features in the lexicon "count differently" and that it provides redundant information for inflection.

<sup>6</sup> There are certain points of agreement here with "minimalist morphology"; cf. Wunderlich – Fabri (1995). For differences cf. Wurzel (in prep.).

*Wurzel* 'root' – plural *die Amseln, Inseln, Wurzeln*. By contrast, non-feminines with the same phonological form usually inflect according to the strong paradigm: *der Beutel* 'bag', *Vogel* 'bird', *das Segel* 'sail' – plural *die Beutel, Vögel, Segel*. In modern German, some originally feminine nouns also appear as masculines or neuters, such as *Kartoffel* 'potato', *Trüffel* 'truffle', or *Partikel* 'particle'. Accordingly, they derive their inflectional forms 'automatically' according to the strong paradigm: *der Kartoffel, der Trüffel, das Partikel* – plural *die Kartoffel, Trüffel, Partikel*.

The strength of speakers' efforts to exploit independent lexical properties in determining inflectional class is seen in the following example, which is related to a semantic property:

**Example 3: Transfer of feminines from the *i*- and *ō*-declensions to the consonantal declension in Proto-Norse**

In Proto-Germanic, feminine nouns in the consonantal, *i*- and *ō*-declensions were distinguished formally in their base forms, e.g., \**vik-z* 'bay', \**leið-iz* 'path' and \**man-ō* 'mane'. Then in Old Icelandic, the base forms of the three inflectional classes had already fallen together formally due to phonological neutralization. The words of all three classes now ended in a consonant, e.g., OIce. *vik*, *leið* and *mon*, thereby losing the phonological indicators for determining their inflectional class (-*z* vs. -*iz* vs. -*ō*). In this area as well, various class transfers may be observed. Significantly in this context, certain words from the *i*-declension like *ǫnd* 'duck' and *ǫlpt* 'swan' and words from the *ō*-declension like *hind* 'hind (of venison)' transferred to the consonantal declension more or less consistently already in Old Icelandic.

The *i*-declension is the largest of the three inflectional classes involved and includes by far the greatest number of nouns ending in a consonant. Thus, according to our previous findings, we would expect to see transfers from the other two classes into this one. Nevertheless, words transfer from this class and from the large *ō*-class into the small consonantal class. This seems surprising at first glance. When one examines the facts more closely, however, it becomes clear that the class of consonantal feminines contains, among others, a large collection of animal names, e.g., *geit* 'goat', *gōs* 'goose', *lús* 'louse', *mús* 'mouse' and *tík* 'bitch', while each of the other two classes contains only a few animal names. The interpretation of these class transfers has to do with speakers' taking the semantic property 'animal' as a criterion for the class membership of lexemes. Due to the purely random division of words across the three classes, the consonantal declension became the preferred class for feminine an-

imals ending in a consonant, and appropriate words of both competing classes joined this one. Thus speakers exploit not only the common, fundamental syntactic property of gender and the equally common, fundamental phonological property of word ending, but also the quite specific semantic property 'animal' for class specification, a property which (as far as anyone knows) had never played a role in Germanic and Nordic grammar before.<sup>7</sup> Neither the synchronic grammar before the transfers nor the synchronic grammar after the transfers shows that this was the case. The use of the semantic property 'animal' in morphological classification becomes visible only from an analysis of language change, that is, from a diachronic analysis.

Let us summarize the results of this section in the following hypothesis:

- (D) The inflectional class membership of the lexemes tends to be determined by independent properties, specifically by phonological properties of the base form or by the syntactic and semantic properties of the lexeme.

#### 4. The specification of marked inflectional class membership

We have already seen with the transfer of weak to strong masculine nouns in German that certain language changes may only be explained adequately if one assumes that unmarked class membership is not explicitly specified in the lexicon, while marked class membership is. But this says nothing about how a marked inflectional class should be represented in the lexicon. It is frequently assumed that the inflectional behaviour of words is fixed in the lexicon simply by a direct specification of inflectional class, i.e., by diacritic features of the type [CS X] (as we did above for simplicity's sake). Yet consider the following example:

##### **Example 4: The gradual transfer of strong to weak verbs in German<sup>8</sup>**

It is well known that strong verbs have been shifting to the weak class since Middle High German. It is often the case that not all forms of a previously strong verb are replaced by weak forms at the same time. Much more often the process happens gradually. At the present there are, among others, two types of verbs that were originally completely strong verbs, which show only a partial replacing of the strong forms by weak forms. One type has a weak

<sup>7</sup> In other languages, this semantic property definitely plays a role in inflectional class membership, as for example in the Bantu languages; cf. Swahili, in which most if not all animal names belong to the *n*-class (along with other nouns).

<sup>8</sup> For details cf. Bittner (1996, *passim*).

present (without vowel change in the 2/3sg.ind.) and a weak preterite with a continued strong past participle and is evidenced by verbs like:

- (5) (a) *salzen* 'to salt' – *er salzt* – *er salzte* vs. *gesalzen*  
 (*sälzt* in the present and *sielz* in the preterite disappeared centuries ago; *gesalzt* has not taken over yet)
- (b) *melken* 'to milk' – *er melkt* – *er melkte* vs. *gemolken*  
 (*milkt* in the present and *molk* in the preterite are archaic; the participle *\*gemelkt* does not (yet) exist)

The other type has a weak present with a continued strong preterite and strong perfect participle and is seen in verbs like:

- (6) (a) *gebären* 'to bear a child' – *sie gebärt* vs. *sie gebar* – *geboren*  
 (*gebiert* in the present is archaic; *\*gebärte* in the preterite and the participle *\*gebärt* do not (yet) exist)
- (b) *gären* 'to ferment' – *es gärt* vs. *es gor* – *gegoren*  
 (*gärte* in the preterite has not won out yet; the participle *\*gegärt* does not (yet) exist).

That in this kind of partial transfer transitional stages of language development need not necessarily occur only in the short-term is demonstrated by the verb *mahlen* 'to grind', for which the inflectional forms *er mahlt* – *er mahlte* but *gemahlen* have been standardized.

Such changes (which are not rare occurrences) are not class transfers but adaptations of some of the inflectional forms to the forms of another class, whereby "mixed inflections" are created. They provide evidence that the words in the lexicon are not specified according to inflectional class, but according to the occurrence of certain inflectional forms in their paradigms. One could assume then that all words that have an inflection that differs from the default case, that is, words that have marked inflection, show further inflectional forms as definitive in the lexicon in addition to the base forms. From these additional forms, it is clear how to construct the rest of the paradigm. Thus weak non-feminine nouns ending in a consonant like *Mensch* 'human being' and *Bär* 'bear', for example, would have to be specified according to their respective plural forms *Menschen* and *Bären* in addition to their base forms. The rest of the inflectional paradigm follows predictably from the plural form. But this assumption is not without its problems, as example 5 shows:

**Example 5: Transfer of strong masculine nouns to the umlaut class in German**  
 Since Middle High German, strong masculine nouns with *e*- and  $\emptyset$ -plurals that originally formed their plurals without umlaut have gradually been adopting the umlaut-plural. Older cases of *e*-plurals that completed this class transfer during previous stages of the German language include *Baum* ‘tree’, *Hut* ‘hat’, *Klang* ‘sound’, and *Schatz* ‘treasure’, with the new plural forms *die Bäume*, *Hüte*, *Klänge*, and *Schätze*. In this century words like *Mops* ‘pug dog’ *Rumpf* ‘torso’ and *Strand* ‘beach’, among others, have switched, cf. *die Möpse*, *Rümpfe*, *Strände*. Among the nouns with  $\emptyset$ -plurals, words like *Garten* ‘garden’, *Ofen* ‘oven’, *Vater* ‘father’, and *Vogel* ‘bird’ have adopted the umlaut-plural, and recently words like *Bogen* ‘curve, bow’ and *Kasten* ‘box’ as well, cf. *die Gärten*, *Öfen*, *Väter*, and *Vögel*, similarly *die Bögen*, *Kästen*. From a large inventory of other such nouns, some like *Hammel* ‘castrated ram’, *Hammer* ‘hammer’ and *Wagen* ‘car’ now take on unumlauted plural forms (especially in South German/Austrian) in addition to the plurals without umlaut, thus *die Hämmel*, *Hämmer*, *Wägen*.<sup>9</sup>

In cases like this, it is important to note that the plural forms of the lexemes are not replaced as a whole by new forms, but rather a new, additional marker, umlaut, is introduced into the existing plural forms. What is marked before the change is the absence of umlaut in the plural forms, but not the plural form as a whole, because the *e*- and the  $\emptyset$ -plurals are unmarked for these nouns. Marked (but not suppletive) inflection is evidenced, at least in such cases, not by the appearance of marked inflectional forms, but by the appearance of marked markers or by the absence of unmarked markers. If one represented the complete plural form in the lexicon, this fact would be neglected.

This results in the assumption that what is specified in the lexicon are not complete marked inflectional forms, but marked markers or, as in this case, the absence of unmarked markers. Therefore, a noun like *Mops* has the inflectional specification [ $\emptyset$ Uml/PL] ‘no umlaut in the plural’ before the class transfer, which specification is removed by the change from *Mops* – *Mopse* to *Mops* – *Möpfe*. Based on this model, if one makes the not implausible generalization that inflectional specifications are represented in the lexicon of a language uniformly, then one is forced to assume that in languages like German the specification of inflection follows generally by reference to the markers. Let me add that it is highly compatible with the “minimalist” hypothesis that the

<sup>9</sup> Incidentally, the plural forms *Hämmer*, *Hämmel* and *Wägen* have been sanctioned by the *Duden*, cf. DUDEN (1996).

lexical entries of words contain no specifications of predictable properties. If one included the complete plural form of strong masculines without unlauded plurals, for example *Hunde*, the plural of *Hund* 'dog', then the stem, which has a completely predictable form in the plural, would be represented in the lexicon twice.

At this point, let us formulate a final hypothesis:

- (E) The specification of inflectional class in the marked cases is not guided by a diacritical specification of the inflectional class, but by reference to concrete inflectional properties of the lexemes. What is specified in the lexicon are category markers (in German: suffixes or vowel alternations) as indicators of the inflectional paradigm.

### 5. Summary

It has been my purpose to demonstrate how, in the area of morphological lexical representations based on an analysis of concrete language changes, one may derive well-founded hypotheses about how speakers organize their grammar, that is, as to how the synchronic grammar is structured. The hypotheses presented here (and others that I was not able to discuss in this short paper) do not yet form a theory of morphological lexical representation. First, the whole area has not been considered in all of its structural aspects; several "holes" remain to be filled in by future studies. Second, the results, which at this point are based only on facts from the Germanic and similarly structured languages, should be applied to language changes from other languages that are structured differently. Thus (who would deny it), there still remains much to be done on the way to a diachronically adequate theory of morphological lexical representation as part of a corresponding morphological theory. Nevertheless, one can say that the methodological approach chosen here, the path from change to structure, has proved fruitful.

### References

- Bittner, A. 1996. *Starke 'schwache' Verben – schwache 'starke' Verben. Deutsche Verbflexion und Natürlichkeit.* Stauffenburg, Tübingen.
- DUDEN. 1996. *Duden. Rechtschreibung der deutschen Sprache*, 21. Auflage. Dudenverlag, Mannheim, Leipzig, Wien, Zürich.
- Kiparsky, P. 1968. Linguistic universals and linguistic change. In: Bach, E.–Harms, R. (eds): *Universals in linguistic theory*, 171–202. Holt, Rinehart and Winston, New York.

- Vennemann, T. 1989. Language change as language improvement. In: Orioloes, V. (ed.): *Modelli esplicativi della diacronia linguistica*, 11–35 Giardini Editori, Pisa.
- Wunderlich, D. – Fabri, R. 1995. Minimalist morphology. In: *Zeitschrift für Sprachwissenschaft* 14: 236–94.
- Wurzel, W.U. 1994. *Grammatisch initiiertes Wandel*. Unter Mitarbeit von A. Bittner und D. Bittner. Projekt "Prinzipien des Sprachwandels". Brockmeyer, Bochum.
- Wurzel, W.U. In prep. *Von der Diachronie zur Synchronie: Zur Struktur von Lexikonrepräsentationen, Flexionsformen, Flexionsparadigmen und Flexionssystemen*.

Address of the author: Wolfgang Ullrich Wurzel  
Zentrum für Allgemeine Sprachwissenschaft,  
Typologie und Universalienforschung  
Jägerstrasse 10/11  
D-10117 Berlin  
Germany  
wurzel@zas.gwz-berlin.de



NACHRUF

CLAUS JÜRGEN HUTTERER  
(2.11.1930–17.12.1997)

Kurz vor Weihnachten erreichte die ungarische Wissenschaftlichkeit die Nachricht, daß Dr. Dr. h. c. Claus Jürgen Hutterer, ordentlicher Professor für Germanistik an der Universität Graz, verstorben ist.

Sein Tod ist ein großer Verlust für die Germanistik und ungarländische Dialektologie, sowie für die Sprachinselforschung.

Claus Jürgen Hutterer wurde am 2.11.1930. in Budapest geboren, sein Vater stammte aus dem ungarndeutschen Dorf Geschitz/Várgesztes im Schildgebirge/Vértés. 1949 begann er sein Studium an der Budapester Loránd Eötvös Universität, wo er Hungaristik und Indogermanistik studierte. 1953 hat er als Lehramtskandidat das Magisterium erworben. Zwischen 1954 und 1958 hat er im Auftrag der Ungarischen Akademie der Wissenschaften das Germanistikstudium—vor allem auf dem Gebiet der deutschen/ungarndeutschen Dialektologie—unter der Betreuung von dem besten Vertreter der deutschen Sprachinselforschung, Viktor Schirmunskij fortgesetzt und die Laut- und Wortgeographie der deutschen Mundarten in Mittelungarn bearbeitet. Mit der entsprechenden Dissertation wurde ihm der Titel des Kandidaten der philosophischen Wissenschaften verliehen. (*Das Ungarische Mittelgebirge als Sprachraum*, Md. Studien 24, Halle/Saale). Von 1958–1985 war er Mitarbeiter der Philosophischen Fakultät der Eötvös Universität, bis 1968 war er als Oberassistent, zwischen 1968–70 als Dozent und ab 1970 als ordentlicher Universitätsprofessor tätig.

In Österreich hat er aufgrund einer Berufung zwischen 1972–75 als Gast- und Honorarprofessor, ab 1975 als ordentlicher Universitätsprofessor an der Universität Graz gearbeitet. 1968 erlangte er an der Ungarischen Akademie der Wissenschaften mit seiner Arbeit über die germanischen Sprachen (*Die germanischen Sprachen*—kurz vor seinem Tode bearbeitete er die 4. Auflage) den Titel Doktor der Wissenschaften.

1996 hat ihm die Budapester Loránd Eötvös Universität in Würdigung seiner Leistungen auf dem Gebiet der Germanistik, der allgemeinen Sprachwissenschaft, besonders der Sprachinselforschung und der ungarndeutschen Dialektologie und Sprachsoziologie das Ehrendoktorat (doctor et professor philo-

sophiae honoris causa) auf Vorschlag der Philosophischen Fakultät, verliehen. Von 1962 bis 1982 war er Moderator und Mitherausgeber der international anerkannten linguistischen Zeitschrift der Ungarischen Akademie der Wissenschaften *Acta Linguistica*. Durch diese Tätigkeit förderte er zahlreiche junge angehende Anglisten und Germanisten auf ihrer wissenschaftlichen Laufbahn. In Österreich wirkte er auch als Mitherausgeber der Reihe *Schriften zur deutschen Sprache in Österreich* mit. Seine Bücher, Studien, wissenschaftlichen Beiträge und Rezensionen erschienen ab Mitte der 1950er Jahre in deutscher, englischer, französischer, spanischer, russischer und holländischer Sprache in Ungarn, Deutschland, Österreich, Rußland, in den USA, Uruguay, Holland, Frankreich und Italien. An der Grazer Universität hat er für zwei Perioden auch die Aufgaben des Institutsvorstandes wahrgenommen.

Hutterer hat die deutsche Mundartforschung in Ungarn viel zu verdanken. In einer Zeit, in der es nicht unbedingt lobenswert war sich mit Fragen der deutschen Minderheit—wissenschaftlich—zu beschäftigen, hielt er an der Budapester Eötvös Universität, im Deutschen Seminar seine Lehrveranstaltungen über die deutschen Mundarten in Ungarn, versuchte seine Schüler auf bisher wenig bearbeitete, weniger bekannte ungarndeutsche Mundartgebiete aufmerksam zu machen, erarbeitete die Grundsätze und Fragebücher zum Ungarndeutschen Sprachatlas sowie zum Ungarndeutschen Wörterbuch, gab Anleitungen zur Schaffung eines Tonarchivs der deutschen Mundarten in Ungarn, trieb selbst Feldforschung mit seinen Studenten, und hielt daneben die Hauptvorlesung sowie die Oberseminare für alle Germanistik- und Anglistik-Studenten über die germanischen Sprachen. Er lenkte die Aufmerksamkeit auf die Untersuchung der soziologischen Staffelung der deutschen Mundarten in Ungarn, und so ist es auch selbstverständlich, daß das Zigeunerische und Jiddische in Ungarn und Europa auch zu seinen Forschungs- und Seminarthemen gehörten.

Grundlegendes schuf Hutterer mit der historischen Laut- und Wortgeographie der deutschen Mundarten im Ungarischen Mittelgebirge. Diese präzise, mehr als 150 Forschungspunkte umfassende sprachgeographische und sprachsoziologische Arbeit, die auch eine Wortbildungslehre beinhaltet, war für die Sprachinseldialektologie sowie die allgemeine Sprachwissenschaft von weitreichender Bedeutung. Die Analyse der Mundarten brachte ihn zur Formulierung der Sieben Thesen zur Dialektforschung, von denen besonders die These 2—*“Die Mundart ist demnach ein in sich faßbarer auch nach ihrem eigenen System erklärbarer Komplex, mit einem strukturalistischen Terminus: ein eigenständiges Korpus, dessen Erforschung letzten Endes auch methodologisch*

*der Erforschung sonstiger Korpora ähnlicher Rangordnung gleichgesetzt werden kann.*—aufgrund seiner tiefbohrenden Forschung erstellt werden konnte.

Hutterer hat in den 60er Jahren die Grundprinzipien des Ungarndeutschen Sprachatlasses und des Wörterbuchs der Ungarndeutschen Mundarten ausgearbeitet, und eine großangelegte Materialsammlung in mehr als 500 Forschungspunkten veranlaßt. Seinen Anleitungen nach liegt das Grundmaterial zum Atlas aus Südungarn bereits gesammelt vor. Auch die Budapester Skandinavistik (heute mit einem selbständigen Lehrstuhl) verdankt seine Gründung Hutterer. Vor seiner Grazer Berufung war er in Budapest Vorstand der Fachgruppe für Allgemeine Germanistik (Skandinavistik). Seine in ungarischer Sprache abgefaßte *Gotische Grammatik* (Budapest 1974) wollte er in diesem Jahr in deutscher Sprache veröffentlichen.

In den 90er Jahren interessierten ihn Fragen der Konvergenz und Divergenz in der Volkskultur der Deutschen im Karpatenbecken und die Sprachentwicklung anderer ethnischer Minderheiten. Das Jiddisch in Ungarn sowie das wallachisch-zigeunerische waren auch wichtige Themen seiner Forschungen. Seinen Meister ehrte er mit der Herausgabe Viktor Schirmunskij-s *Linguistische und ethnographische Studien 1926–1931* (München 1992). Zu seinem 60. Geburtstag wurden von der Budapester Germanistischem Institut im Band 6. der Ungarndeutschen Studien seine wichtigsten Aufsätze herausgegeben (*Aufsätze zur deutschen Dialektologie*).

Von Graz aus, wo er als Ordinarius tätig war und auch die steirischen Mundarten erforschte, hielt er den regelmäßigen Kontakt zur Budapester Germanistik aufrecht, empfing ständig Forscher und Studenten aus Ungarn zu Studienaufenthalten und beteiligte sich an der Arbeit am Sprachatlas und dem Mundartwörterbuch im Germanistischen Institut. Seit Anlauf des akkreditierten Germanistikprogramms wirkte er auch als stundengebender Professor in der PhD-Ausbildung des Germanistischen Instituts mit, verbrachte jährlich mehrmals längere Zeit in Ungarn als Gastprofessor, und hielt Seminare ab. Claus Jürgen Hutterer trennte sich nie von Ungarn, hielt stets enge Beziehungen zur Budapester Universität aufrecht und war durch die von ihm gegründete Dialektologie-Schule stets im wissenschaftlichen Leben der Philosophischen Fakultät der Eötvös Universität präsent.

Mit Claus Jürgen Hutterer verlieren seine Schüler, Freunde und Mitarbeiter einen offenen, aufrichtigen Menschen, der immer bereit war zu helfen. Als sein Schüler erinnere ich mich an die Budapester Jahre, in denen er in seinem Arbeitszimmer im Zigarettenschmuck in langen Gesprächen und Diskussionen die Grundlagen der Skandinavistik, des ungarndeutschen Sprachatlasses sowie des Wörterbuchs des Jiddischen und Zigeunerischen in Ungarn, des Tonarchi-

ves und der Fragebücher umriß. Sein Lieblingsforschungspunkt war Deutschpilsen/Nagybörzsöny. Unvergesslich bleibt uns seine Arbeitsmethode bei der Sammlung des Sprachmaterials oder bei Tonaufnahmen. Er führte mit den Gewährsleuten in der Mundart Gespräche sowohl in Deutschpilsen/Nagybörzsöny, als auch in der Zips, in Westungarn oder in der Batschka. Es gab keine mundartlichen Ausdrücke die er nicht gekannt hätte.

Wir wußten, daß er krank war, aber glaubten es nicht. Einige Tage vor seinem Tod planten wir noch die Gesamtausgabe der Schriften von Eugen Bonomi, dem Erforscher des Ofner Berglandes und die Kontrolle der Atlaskarten. Nun müssen diese Arbeiten ohne Ihn vollendet werden. Er wird uns allen sehr fehlen.

*Karl Manherz*

# ACTA LINGUISTICA HUNGARICA

VOLUME 47, NUMBERS 1–4, 2000

---

## CONTENTS

|                                                                                                                                     |     |
|-------------------------------------------------------------------------------------------------------------------------------------|-----|
| Preface .....                                                                                                                       | 1   |
| <i>Bartos, H.</i> : VP-ellipsis and verbal inflection in Hungarian .....                                                            | 3   |
| <i>Bisetto, A.–Scalise, S.</i> : Complement selection in morphology and syntax .....                                                | 25  |
| <i>Cetnarowska, B.</i> : Resultative adjectives in Polish .....                                                                     | 47  |
| <i>Di Sciullo, A.-M.</i> : Asymmetries: Consequences for morphological configurations and paradigms .....                           | 81  |
| <i>Dressler, W.U.–Ladányi, M.</i> : Productivity in word formation (WF): A morphological approach .....                             | 103 |
| <i>Elekfi, L.</i> : Semantic differences of suffixal alternates in Hungarian .....                                                  | 145 |
| <i>Español-Echevarría, M.–Ralli, A.</i> : Case mismatches in Greek: Evidence for the autonomy of morphology .....                   | 179 |
| <i>Gaeta, L.</i> : On the interaction between morphology and semantics: the Italian suffix -ATA .....                               | 205 |
| <i>Gràcia, L.–Turon, L.</i> : On appreciative suffixes .....                                                                        | 231 |
| <i>Ortmann, A.</i> : Where plural refuses to agree: feature unification and morphological economy .....                             | 249 |
| <i>Pirrelli, V.–Battista, M.</i> : On the interaction of paradigmatic and syntagmatic stem alternation in Italian conjugation ..... | 289 |
| <i>Sleeman, P.–Verheugd, E.</i> : Deverbal adjectivalization as a gradual process .....                                             | 315 |
| <i>Spencer, A.</i> : Inflection and the lexeme .....                                                                                | 335 |
| <i>Wurzel, W.U.</i> : Morphological properties in the lexicon: Diachronic evidence .....                                            | 345 |
| Nachruf: Claus Jürgen Hutterer (2.11.1930–17.12.1997) .....                                                                         | 357 |

Printed in Hungary  
PXP Ltd., Budapest

## INFORMATION FOR CONTRIBUTORS

Please provide typewritten double-spaced manuscripts with wide margins (50 characters per line, 25 lines per page). Type on one side of the paper only and number all the pages consecutively. The following information should be provided on the first page: the title of the article (or, in case of reviews, the title of the book under discussion, the name of the publisher, the place and date of publication as well as the number of pages); the author's full name and address and an abbreviated title of the paper (not exceeding 45 characters including spaces). Only original papers will be published and a copy of the Publishing Agreement will be sent to the authors of papers accepted for publication. Manuscripts will be processed only after receiving the signed copy of the agreement. Authors are requested to send two hard copies of their manuscript + a floppy disk under DOS.

The manuscript should be accompanied by an abstract of about 100 words. It should be typed on a separate sheet of paper. Tables, diagrams and illustrations (with the author's name and an Arabic number) should be presented on separate sheets. Captions should be typed on a separate sheet and placed at the end of the manuscript. Footnotes should be typed double-spaced on a separate sheet at the end of the article. All language examples (and only these) are to be italicized (single underlining). Citations in the text should be enclosed in double quotation marks (" ") in case of a paper written in English, „ " in German and « » in French).

Reference to a publication should be made by the name of the author, the year of publication and, when necessary, by page numbers in the following ways:

- ... as described by Schmidt (1967) . . .
- ... as referred to by Hoover (1967, 56–78; 1976, 43).
- ... mentioned by several authors (Elgar 1978, 77; Williams 1981, 154–6) . . .

An alphabetically arranged list of references should be presented at the end of the article as follows:

- Bárczi, G. 1958a. *Magyar hangtörténet* [The history of Hungarian sounds]. Akadémiai Kiadó, Budapest.
- Bárczi, G. 1958b. *A szótövek* [Word stems]. Akadémiai Kiadó, Budapest.
- Lakoff, G.–Peters, P.S. 1969. Phrasal conjunction and symmetric predicates. In: Reibel, D.–Schane, S. (eds): *Modern studies in English*, 113–41. Prentice-Hall, Englewood Cliffs.
- Leben, W.R. 1980. A metrical analysis of length. In: *Linguistic Inquiry* 11: 497–509.
- Ross, J.R. 1967. *Constraints on variables in syntax*. Ph.D. dissertation. MIT, Cambridge MA.

For marking subsections decimal notation should be applied. Do not use more than four digits if possible.

Examples within the text should be marked in italics. Meanings are to be rendered between inverted commas (' '). If glosses are given morpheme by morpheme, the initial letter of the gloss should be placed exactly below that of the example. Grammatical morphemes can be abbreviated in small case letters connected to the stem or the other morphemes by a hyphen. No period should be applied in such abbreviation. For example:

- (1) (a) A sólymaid            elszálltak  
          the falcon-gen-pl-2sg away-flew-3pl  
          'Your falcons have flown away.'

Examples can be referred to in the text as (1a), (1a–d), etc.

One proof will be sent to the author. Please read it carefully and return it by air mail to the editor within one week.

# ACTA LINGUISTICA HUNGARICA

VOLUME 47, NUMBERS 1-4, 2000

---

## CONTENTS

|                                                                                                                                             |     |
|---------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Preface.....                                                                                                                                | 1   |
| <i>Bartos, H.</i> : VP-ellipsis and verbal inflection in Hungarian.....                                                                     | 3   |
| <i>Bisetto, A.</i> - <i>Scalise, S.</i> : Complement selection in morphology and syntax                                                     | 25  |
| <i>Cetnarowska, B.</i> : Resultative adjectives in Polish.....                                                                              | 47  |
| <i>Di Sciullo, A.-M.</i> : Asymmetries: Consequences for morphological configurations and paradigms.....                                    | 81  |
| <i>Dressler, W.U.</i> - <i>Ladányi, M.</i> : Productivity in word formation (WF): A morphological approach.....                             | 103 |
| <i>Elekfi, L.</i> : Semantic differences of suffixal alternates in Hungarian.....                                                           | 145 |
| <i>Español-Echevarría, M.</i> - <i>Ralli, A.</i> : Case mismatches in Greek: Evidence for the autonomy of morphology.....                   | 179 |
| <i>Gaeta, L.</i> : On the interaction between morphology and semantics: the Italian suffix -ATA.....                                        | 205 |
| <i>Gràcia, L.</i> - <i>Turon, L.</i> : On appreciative suffixes.....                                                                        | 231 |
| <i>Ortmann, A.</i> : Where plural refuses to agree: feature unification and morphological economy.....                                      | 249 |
| <i>Pirrelli, V.</i> - <i>Battista, M.</i> : On the interaction of paradigmatic and syntagmatic stem alternation in Italian conjugation..... | 289 |
| <i>Sleeman, P.</i> - <i>Verheugd, E.</i> : Deverbal adjectivalization as a gradual process.....                                             | 315 |
| <i>Spencer, A.</i> : Inflection and the lexeme.....                                                                                         | 335 |
| <i>Wurzel, W.U.</i> : Morphological properties in the lexicon: Diachronic evidence.....                                                     | 345 |
| Nachruf: Claus Jürgen Hutterer (2.11.1930-17.12.1997).....                                                                                  | 357 |



1216-8076(2000)47:1/4;1-O