Acta Linguistica Hungarica

AN INTERNATIONAL JOURNAL OF LINGUISTICS

Volume 42, Numbers 1–2, 1994

Akadémiai Kiadó, Budapest

Acta Linguist. Hung. HU ISSN 1216-8076

ACTA LINGUISTICA HUNGARICA

AN INTERNATIONAL JOURNAL OF LINGUISTICS

Acta Linguistica Hungarica publishes papers on the subjects of Finno-Ugric, Slavonic, Germanic, Oriental and Romance linguistics as well as general linguistics in English, German and French.

Acta Linguistica Hungarica is published in yearly volumes of four issues by

AKADÉMIAI KIADÓ

Publishing House of the Hungarian Academy of Sciences

H-1117 Budapest, Prielle Kornélia u. 19–35

Manuscripts should be addressed to:

ACTA LINGUISTICA HUNGARICA
H-1014 Budapest, Színház u. 5–7

Correspondence with the editors and publishers should be sent to the same address.

Orders may be placed with

AKADÉMIAI KIADÓ
H-1519 Budapest, P.O. Box 245

Subscription price for Volume 42 (1994) in 4 issues US$ 84.00, including normal postage, airmail delivery US$ 20.00.

Acta Linguistica Hungarica is abstracted/indexed in Current Content–Arts and Humanities, Arts and Humanitites Citation Index and Bibliographie Linguistique/Linguistic Bibliography

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). The contribution proper should begin on page 2.

CONTENTS

Book reviews

PRINTED IN HUNGARY
AKADÉMIAI KIADÓ ÉS NYOMDA, BUDAPEST

GUEST EDITORS' NOTE

During the past ten years Hungarian speech research has become internationally known via its representatives' participation at diverse conferences, in various publications and by lectures that researchers have been giving abroad. However, this is the first time that a collection of papers specifically discusses, and presents the results of, investigations approaching speech from several different angles.

The background to this issue of Acta Linguistica Hungarica was a national conference on "Theoretical and applied phonetics" which was held in 1992 in Budapest, organized by the Phonetics Laboratory of the Research Institute for Linguistics. The aim of that conference was to show where Hungarian speech research stands just before the end of this century. More than twenty papers had been read at the conference from which eight were selected for inclusion in this collection. The only criterion for selection was that the papers should represent a variety of theoretical and applied investigations concerning speech. According to this aim the present eight papers cover wide-apart points in the field of speech research from phonology up to speech disorders. These papers do not have any points of contact among themselves except that their authors sometimes co-work on studies of an interdisciplinary character. On the other hand, this collection of articles highlights some puzzling topics of current speech research in Hungary.

One of the theoretical papers within the area of phonology is Péter Siptár's "Palatalization rules in Hungarian" which describes Hungarian palatalization processes at there different depths applying a Lexical Phonology framework and autosegmental formalism. Mária Gósy's "The effect of syntagmatic context on lexical access" makes an attempt to raise some new points concerning the problem of lexical access based on experimental data in the context of a non-connectionist approach.

Although quantitative measurements have a long tradition in phonetics, the issue of the temporal organization of speech seems to require some revision from time to time either because new techniques provide new possibilities to search 'hidden' micro- or macrodata or, because natural speech is supposed to change as the decades go by. Gábor Olaszy's paper contains important language-specific results about the durational changes in spontaneous speech using a data-base and a modern speech processing system while Ilona Kassai's work aims to substantiate the claim that speech tempo has accelerated in Hungarian during the past 100–120 years.

Tamás Szende's "Do historical changes repeat themselves?" compares the historical tendency of 'Two-Open-Syllable Shortening', also known as 'Horger's Law', with a present-day postlexical lenition process referred to as 'Fast-Speech Syllable Elision'.

The remaining three papers show a gradual swerve from theories to applications and practice. "Voice quality changes in old age" by Boglárka Balázs fits into the research trend of 'gerontology' analyzing speech productions of elderly people. Klára Vicsi with her co-authors describes a new computer based teaching and training system for speech handicapped persons to control their articulation by visual feedback. Ágnes Hegyi in her paper "Cognitive language therapy in global aphasia" discusses results based on a new therapy as an integration of medical, pedagogical, psychological, and linguistic approaches.

The present issue also includes reviews of some recently published works that share the same topics, i.e. speech research from theories to applications.

Budapest, 1994.

Mária Gósy *Péter Siptár*

Acta Linguistica Hungarica, Vol. 42 (1–2), pp. 5–32 (1994)

PALATALIZATION RULES IN HUNGARIAN*

PÉTER SIPTÁR

1. Introduction

In this paper, the term 'palatalization' is meant to refer to a phonological process whereby a palatal consonant, /j/, /t^y/, /d^y/, or /n^y/, affects a preceding (dental) consonant such that the latter becomes (palato-alveolar or) palatal.[1]

Hungarian phonology has palatalization processes of three different depths: the morphophonological palatalization of *t*-final verbs in the imperative (*üt/üss* 'hit', *önt/önts* 'pour'), lexical palatalization as in *látja* 'he sees it', *hordja* 'he carries it', *fonja* 'he braids it', *falja* 'he devours it', and postlexical palatalization as in *átjáró* 'passage', *védjegy* 'trade mark', *van joga* 'he's got the right (to)', *feljön* 'he comes up'. The most important difference between the last type and the other two (apart from optionality) is that the output is not fused with the trigger (cf. *átjáró* [a:tja:ro:] ~ [a:t^yja:ro:], *[a:t^y:a:ro:]) and that the trigger may be a palatal stop or nasal as well as a /j/: *két nyúl* 'two rabbits', *van gyufa* 'we've got matches'.

This paper will be organized as follows. In section **2**, palatalization processes of all three types will be discussed and formalized both in classical–generative and in autosegmental terms. The geometry assumed is based on the model proposed by Clements (1985); cf. Siptár (1993) for a detailed discussion of that geometry as applied to Hungarian. The distinctive feature analysis of Hungarian consonants is identical to that proposed in Siptár (1993), a revised version of the analysis first suggested in Vago (1980), with a single but crucial exception: /l/ will be taken to be [+cont] rather than [–cont]. That move will be motivated in **2.3** below. Section **3** will explore the implications of the anal-

* This research was supported in part by OTKA [Project T 013863].

[1] In particular, we wish to exclude the fully automatic, low-level, non-neutralizing—and probably non-language-specific—type of 'phonetic palatalization' that is triggered by nonlow front vowels and /j/ and that produces more or less palatalized velars/dentals/labials as e.g. in *kín* 'torture' vs. *kár* 'damage' vs. (labialized) *kút* 'well'. This coarticulatory process will be ignored here.

ysis presented in **2.1–2.3** with respect to the stratal organization of phonology. Finally, section 4 offers a few conclusions with respect to the representational and organizational issues raised in this paper.

2. Data and discussion

2.1. The imperative of *t*-final verbs

Final *t* in verb stems undergoes three kinds of changes before imperative *j*:

– it surfaces as [š] after short vowels as in *üt/üss* 'hit';

– it becomes [č] after sonorant consonants as in *hajt/hajts* 'drive', *önt/önts* 'pour', *olt/olts* 'extinguish'; and

– it deletes after obstruents. That obstruent may be of two sorts: /š/ in *fest* 'paint' and /s/ in verbs of the *oszt* 'divide' type. The Hungarian Explanatory Dictionary includes 117 items of that kind, all of which involve a morpheme boundary between the /s/ and the /t/.

Of V:*t*-final verbs, *lát* 'see', *bocsát* 'let go', and *lót(-fut)* 'rush about' pattern with short vowel stems, whereas *fűt* 'heat', *hűt* 'chill', *műt* 'operate on', *szít* 'incite', *tát* 'open wide', and *vét* 'err', as well as hundreds of verb stems involving causative *-ít*, behave like sonorant + *t*-finals (cf. Jakab 1967). That distribution could be described in three different manners. First, we could rest content with straightforward listing, as we did here. This is the simplest but least satisfactory solution. Second, we could say that the *lát* set is exceptional, and that long vowel stems regularly change their final /t/ into [č]: this is the traditional account (Deme 1961), as well as that suggested by Abondolo (1988, 146). Finally, following an assumption first made by Vago (1980, 72), we could claim that the *lát* set plus short vowel stems constitute the regular (vowel + *t*) class and the *fűt* group should somehow be included in the [č] class. Vago does that by positing an underlying /j/ in the verb stems concerned, one that merges with the preceding vowel (or rather deletes with compensatory lengthening) after the now regular palatalization of /t/ into [č] but before degemination as in *hajts* etc.: /füjt+j/ ⟶ füjč: ⟶ [fü:č:] *fűts* 'heat-IMP' vs. /hajt+j/ ⟶ hajč: ⟶ [hajč] *hajts* 'drive-IMP'. This solution is not as arbitrary as it might seem: the verbs listed behave as if they were /Vjt/-final in a number of other respects as well. For instance, infinitival *-ni* attaches to vowel + *t*-final verbs without an epenthetic (linking) vowel: *ütni* 'to hit', *futni* 'to run', *látni* 'to see', but requires an epenthetic vowel after consonant + *t*-final ones: *osztani* 'to divide', *hajtani* 'to drive', *váltani* 'to change'. The *fűt* set shares the behaviour of the latter group: *fűteni* 'to heat', *hűteni* 'to chill', *szítani* 'to incite', etc.[2]

The only respect in which this account has a touch of arbitrariness is the quality of the assumed underlying segment: why /j/? The reason is that this is the only eligible sonorant that does not occur on the surface in exactly this environment (cf. Vago 1991). Autosegmental (or CV) phonology removes that little arbitrariness, too (cf. Vago 1987, 1989, 1991):

(1)
$$\textit{tát} = \begin{array}{cccc} \mathrm{C} & \mathrm{V} & \mathrm{C} & \mathrm{C} \\ | & \multicolumn{2}{c}{\vee} & | \\ \mathrm{t} & \multicolumn{2}{c}{\mathrm{a}} & \mathrm{t} \end{array} \qquad \textit{lát} = \begin{array}{cccc} \mathrm{C} & \mathrm{V} & \mathrm{V} & \mathrm{C} \\ | & \multicolumn{2}{c}{\vee} & | \\ \mathrm{l} & \multicolumn{2}{c}{\mathrm{a}} & \mathrm{t} \end{array}$$

The foregoing can be summarized as in (2):

(2)
$$\begin{array}{l} \mathrm{t} \longrightarrow \mathrm{\check{s}} \;/\; \mathrm{V}\ \_\_\ \mathrm{j} \\ \mathrm{t} \longrightarrow \mathrm{\check{c}} \;/\; \begin{bmatrix} \mathrm{C} \\ +\mathrm{son} \end{bmatrix} \_\_\ \mathrm{j} \\ \mathrm{t} \longrightarrow \emptyset \;/\; [-\mathrm{son}]\ \_\_\ \mathrm{j} \end{array}$$

In classical generative terms, the three rules can be collapsed as follows (cf. Vago 1980, 71; for a criticism of this collapsed statement of the rule, see Kontra 1992):

(3) *t*-palatalization/deletion

$$\mathrm{t} \longrightarrow \left\{ \begin{array}{c} \left\langle \begin{bmatrix} -\,\mathrm{ant} \\ +\mathrm{strid} \\ +\mathrm{del\ rel} \\ \alpha\ \mathrm{cont} \end{bmatrix} \right\rangle \\ \emptyset \end{array} \right\} \;/\; \left\langle \begin{bmatrix} +\mathrm{son} \\ \alpha\ \mathrm{syl} \end{bmatrix} \right\rangle \_\_ \begin{bmatrix} -\,\mathrm{syl} \\ +\mathrm{son} \\ -\,\mathrm{nas} \\ -\,\mathrm{ant} \end{bmatrix}$$

Now we have to answer the following two questions:

(i) What happens to the /j/?
(ii) How can all this be limited to imperative forms?

[2] Note, however, that *bocsát* 'let go' is assigned to the *fűt* type by the *-ni* test and even has a free variant *bocsájt* to make things worse, yet its imperative has [š] rather than [č].

The above palatalization processes leave the imperative /j/ in a situation where it is preceded by [š], [č], or [s]; what happens to it subsequently is the same as in sibilant-final verbs: /moš/ 'wash' + /j/ ⟶ [moš:], /fut/ 'run' + /j/ ⟶ fuš + j ⟶ [fuš:]. In other words, *t*-palatalization/deletion feeds rule (4); this is illustrated in the derivations in (5):

(4) *j*-assimilation

$$\begin{array}{llcll} [+\text{strid}] & \text{j} & & & \\ 1 & 2 & \longrightarrow & 1 & 1 \end{array}$$

(5)

	/moš+j/	/fut+j/	/hajt+j/	/fešt+j/
t-pal/del (3)	–	š	č	∅
j-assim (4)	š	š	č	š
degemination	–	–	∅	–
other rules	[moš:]	[fuš:]	[hajč]	[feš:]
	'wash!'	'run!'	'drive!'	'paint!'

In standard Hungarian, *t*-palatalization/deletion only applies in imperative forms.[3] This restriction can be indicated in the rule itself; but it would be more appropriate to find some principled explanation. Vago's (1980) solution relies on the claim that both personal and possessive suffixes have non-underlying (epenthetic) *j*'s (they are epenthesized by two different rules, cf. Vago *op. cit.* 68, 105). If the rules of epenthesis are applied after that of *t*-palatalization/deletion (whether by extrinsic ordering or for some other reason), the fact that (3) is restricted to imperative forms is automatically accounted for.[4]

Turning now to an autosegmental formalization of our rules so far, observe first of all that *t*-pal/del is not a possible autosegmental rule. We have to return to the informal statements in (2); on the other hand, in cases covered by (2a) and (2b), an overwhelming majority of all cases, no separate rule of

[3] In a stigmatized version of substandard Hungarian ("*suk-sükölés*"), it also applies before *j*-initial personal suffixes, i.e. in indicative forms as well (cf. Kontra 1992); the /j/ of possessive suffixes, however, does not trigger it in any variety of Hungarian: *botja* *[boš:a] 'his stick', *lantja* *[lanča] 'his lute'.

[4] Possessive *j* is actually quite likely to be epenthetic, cf. Kiefer (1985); and the *j* of verbal personal suffixes may well turn out to be epenthetic, too. The present issue is not whether these segments are underlying or otherwise. However, in section **3** we are going to propose another explanation that does not require either epenthesis rules or extrinsic ordering.

j-assimilation is needed; hence the rule system may become more complicated but the individual derivations will be shorter.[5]

The first thing we have to do is reshuffle (2) to optimize the 'elsewhere' relations involved; we thus obtain a set of three rules each pair of which, (a)/(b), (b)/(c), as well as (a)/(c), is in a counterbleeding relationship:

(6) (a) $\mathrm{t} \longrightarrow \emptyset \;/\; [-\mathrm{son}]\,\underline{\quad}\,\mathrm{j}$

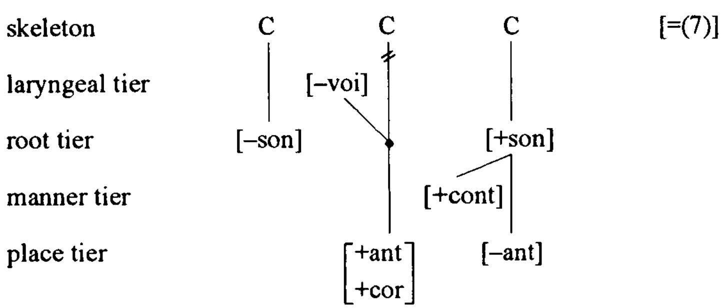
(b) $\mathrm{t} \longrightarrow \check{\mathrm{c}} \;/\; \mathrm{C}\,\underline{\quad}\,\mathrm{j}$

(c) $\mathrm{t} \longrightarrow \check{\mathrm{s}} \;/\; \underline{\quad}\,\mathrm{j}$

Also, we may expand (a) to cover cases like *tetszik* 'it pleases' ⟶ *tessék* 'let it please', *látszik* 'it seems' ⟶ *lássék* 'let it seem' (cf. Olsson 1992, 158):

(7) $\left\{\begin{matrix}\mathrm{t}\\ \mathrm{s}\end{matrix}\right\} \longrightarrow \emptyset \;/\; [-\mathrm{son}]\,\underline{\quad}\,\mathrm{j}$

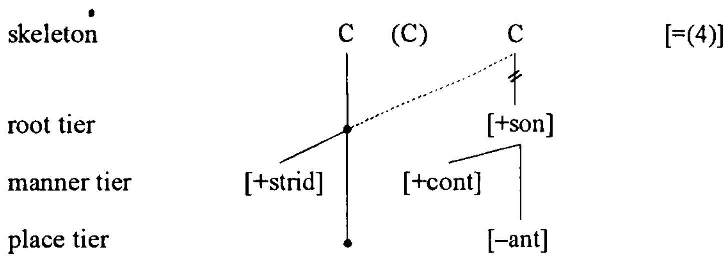
We then get a generalized deletion rule stated in autosegmental terms as in (8):

(8)


This rule will feed (11) below in cases of the *tessék* 'let it please' type; whereas in *ossz* 'divide-Imp', *fess* 'paint-Imp' etc., the output of (8) will undergo (9), the rule that also accounts for cases like *hozz* 'bring-Imp', *moss* 'wash-Imp', *mássz* 'crawl-Imp':

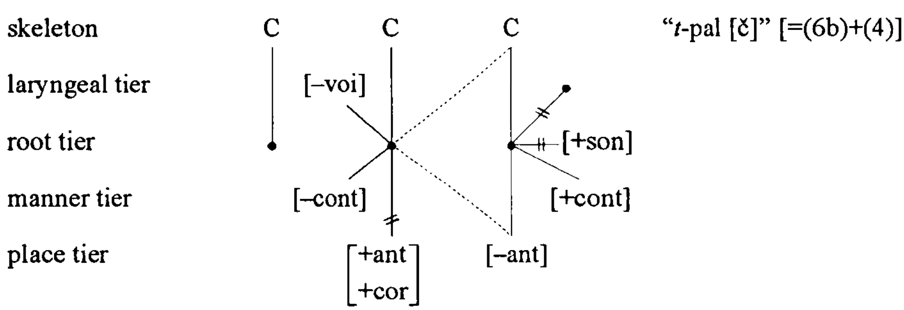
[5] Furthermore, the complication introduced in the rule system is more apparent than real: strictly speaking, (3) is not a single rule. Rather, it is a set of three rules that can be written in a collapsed form, due to abbreviatory conventions of the classical generative framework. See Kontra (1992) for critical discussion.

(9)



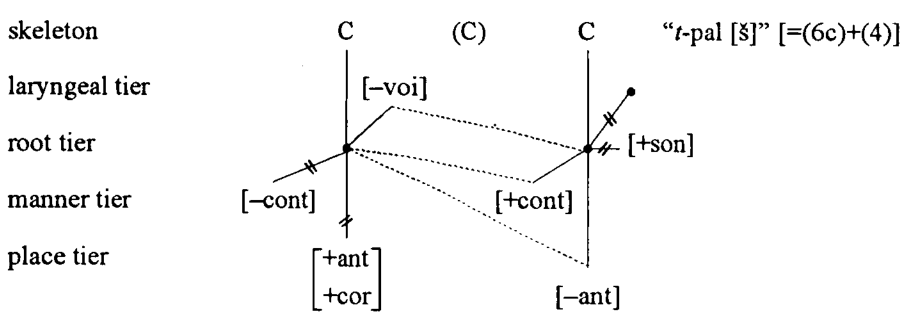
(The parenthesized (C) stands for an optional empty consonantal slot that does not block the application of the rule and remains phonetically unimplemented.) In cases like *önts* 'pour-Imp', *hajts* 'drive-Imp', *fűts* 'heat-Imp', a long affricate is produced that is input to degemination in *önts*, *hajts*, but surfaces as long in *fűts*:

(10)



Finally, in cases like *üss* 'hit-Imp' (and also in cases like *tessék* as we pointed out above), rule (11) applies:

(11)



That is, the /t/ loses its place and manner nodes by delinking and acquires the palatality and continuancy of /j/ by spreading; at the same time, /j/ loses its laryngeal and [son] nodes, becoming a voiceless fricative. The total result is a voiceless palatal geminate fricative, i.e. a [š:].[6]

2.2. Lexical palatalization

The next type of palatalization we will consider (e.g. *látja* [t^y:] 'he sees it', *adja* [d^y:] 'he gives it/let him give it', *kenje* [n^y:] 'let him smear it', *falja* [j:] 'he devours it/let him devour it'; *bátyja* [t^y:] 'his brother', *hagyja* [d^y:] 'he leaves it/let him leave it', *hányja* [n^y:] 'he throws it/let him throw it') is also triggered by /j/ but the result in this case is a true palatal (not palato-alveolar) geminate; the input segments are dental /t d n l/ and—partly vacuously—palatal /t^y d^y n^y/.

Vago's (1980) classical generative solution assumes that /l/ is [−cont], in order to be able to include $l \longrightarrow j$ in the same rule as the other segments, all of which are uncontroversially [−cont]. As observed by Olsson (1992), this solution is problematic with respect to the postlexical behaviour of *lj* sequences. We are going to discuss this point in **2.3**; for the moment let us simply assume that /l/ is [+cont] and its full assimilation to a subsequent /j/ is to be accounted for by a separate rule of *l*-palatalization.[7]

The price we have to pay for characterizing /l/ as [+cont] is that we have to introduce the extra manner feature [lat] to distinguish /l/ from /r/.

[6] Due to Structure Preservation, a lexical phonological rule cannot create "new" types of segments, i.e. ones that do not occur in underlying representations. A postlexical rule of the form (11) would create a [ç:]; since, however, the present rule is a (Level 1) lexical rule, Structure Preservation triggers the relevant default (or redundancy) rule ([−voice, +cont, −ant] $\longrightarrow$ [+strid]), and the output will be [š:].

[7] Vago's solution implies that the palatalization of /l/ produces a long palatal lateral [l^y:] which has to be 'phonetically interpreted' by a $l^y \longrightarrow$ [j] adjustment rule (Vago *op. cit.* 41, (29)), e.g. /tolja/ $\longrightarrow$ tol^yl^ya $\longrightarrow$ [toj:a] 'he pushes it/let him push it'. Notice that Structure Preservation does not help here: then, the output of the lexical palatalization of /l/ could be [j:] but that of the postlexical palatalization of /l/ as in *hol jártál* 'where have you been' should be *[l^y:] (actually, it is also [j:]). Thus, we would need a redundancy rule that remained operative throughout the phonology (including the postlexical component), stating that Hungarian does not have a palatal lateral: if a liquid is [−ant], it must be [−lat] (and, in Vago's system, also [−cont]): $\begin{bmatrix} +\text{son} \\ -\text{nas} \\ -\text{ant} \end{bmatrix} \longrightarrow \begin{bmatrix} -\text{lat} \\ -\text{cont} \end{bmatrix}$.

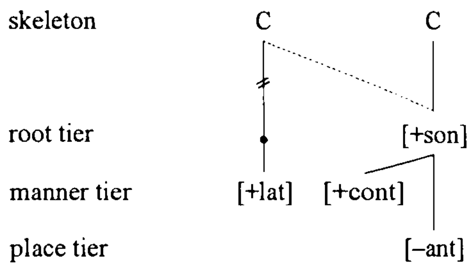
(Alternatively, we could introduce [trill] for /r/.)[8] The rule of *l*-palatalization can then be formulated as (12a) in classical generative terms and as (12b) in the autosegmental framework:

(12) *l*-palatalization

(a)

$$[+\text{lat}] \rightarrow \begin{bmatrix} -\text{ant} \\ -\text{lat} \end{bmatrix} / \;\_\_\; \begin{bmatrix} -\text{syl} \\ +\text{son} \\ -\text{nas} \\ -\text{ant} \end{bmatrix}$$

(b)



The rest of the process can be described by a pair of rules, *à la* (3) and (4), or in a single step, as a rule of mutual assimilation. Consider the standard solution first:

(13) Palatalization

$$\begin{bmatrix} +\text{cor} \\ -\text{cont} \\ -\text{strid} \end{bmatrix} \longrightarrow [-\text{ant}] / \;\_\_\; \begin{bmatrix} -\text{syl} \\ +\text{son} \\ -\text{nas} \\ -\text{ant} \end{bmatrix}$$

(i.e. before /j/, /t d n/ go to [t^y d^y n^y], respectively)

[8] Yet another solution is proposed by Olsson (1992). In his system, /l/ (along with /t d n/) is characterized as [+dental], whereas /r/ (along with /s z t^s d^z/) as [−dental] (both sets being [+ant, +cor], as well as—redundantly—[-lab, -high, -back]). See my review of Olsson (1992) later in this volume.

(14) Palatal Coalescence

$$\begin{bmatrix} -\text{syl} \\ -\text{ant} \\ +\text{cor} \\ -\text{strid} \end{bmatrix} \quad \begin{bmatrix} -\text{syl} \\ +\text{son} \\ -\text{nas} \\ -\text{ant} \end{bmatrix}$$

$$1 \qquad\qquad 2 \qquad\qquad \longrightarrow \quad 1 \quad 1$$

(i.e. $/t^yj/$ goes to $[t^yt^y]$, $/d^yj/$ goes to $[d^yd^y]$, $/n^yj/$ goes to $[n^yn^y]$)

It is possible to collapse (14) with (4), as follows (this is also suggested by Abondolo 1988, 64, although he does not formalize the rule schema):

(15) *j*-assimilation (generalized)

$$\left\{ \begin{matrix} \begin{bmatrix} +\text{cor} \\ -\text{ant} \end{bmatrix} \\ [+\text{strid}] \end{matrix} \right\} \quad \begin{bmatrix} -\text{syl} \\ +\text{son} \\ -\text{nas} \\ -\text{ant} \end{bmatrix}$$

$$1 \qquad\qquad 2 \qquad\qquad \longrightarrow \quad 1 \quad 1$$

(i.e. $/sj/ \longrightarrow [ss]$, $/zj/ \longrightarrow [zz]$, $/t^sj/ \longrightarrow [t^st^s]$, $/d^zj/ \longrightarrow [d^zd^z]$,
$/šj/ \longrightarrow [šš]$, $/žj/ \longrightarrow [žž]$, $/čj/ \longrightarrow [čč]$, $/ǰj/ \longrightarrow [ǰǰ]$,
$/t^yj/ \longrightarrow [t^yt^y]$, $/d^yj/ \longrightarrow [d^yd^y]$, $/n^yj/ \longrightarrow [n^yn^y]$)

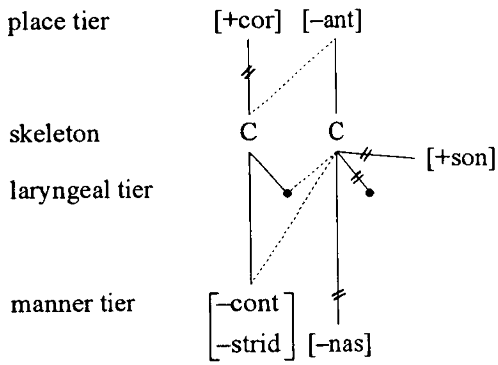
Let us illustrate the analysis so far with the following derivations:

(16)

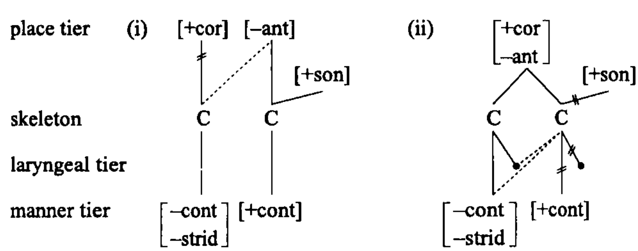
	/hoz+ja/	/la:t+ja/	/had^y+ja/	/du:l+ja/
l-palatalization (12)	–	–	–	j
Palatalization (13)	–	t^y	–	–
j-assimilation (15)	z	t^y	d^y	–
other rules	[hoz:a]	[la:t^y:a]	[had^y:a]	[du:j:a]
	'brings it'	'sees it'	'leaves it'	'destroys it'

Consider now the other solution: replace (13) and (14) by a single rule of mutual assimilation. Although both solutions can be formalized in either framework,[9] the former is more appropriate to the classical generative, while the latter to the autosegmental, formalism. Further differences will be explored in section **3** below.—The rule of lexical palatalization, construed as mutual assimilation, will be formalized as in (17):

(17) Lexical Palatalization



[9] The autosegmental equivalents of (13) and (14) are given in (i) and (ii), respectively; the classical generative equivalent of (17) might be written as in (iii):



(iii) $\begin{bmatrix}\text{+cor}\\ \text{–strid}\end{bmatrix} \quad j$

$1 \quad 2 \rightarrow \begin{bmatrix}1\\ \text{–ant}\end{bmatrix}\begin{bmatrix}1\\ \text{–ant}\end{bmatrix}$

The palatality of /j/ (in the form of [–ant]) spreads leftwards, while the place features of the preceding segment are delinked (in addition to [+cor], the [ant] feature, too, whatever value it might have; the resulting geminate will predictably remain [+cor] since /j/ has that value as well); at the same time, the liquid character of /j/ is removed (by delinking [+son], the manner node with [–nas], and the implicit [+voice] with the laryngeal node) while the laryngeal and manner features of the preceding stop or nasal spread onto it (along with those indicated, [voice] and [nas] as well). Thus, /tj/ and /t^{y}j/ both go to [t^y:], /dj/ and /d^{y}j/ both go to [d^y:], /nj/ and /n^{y}j/ both go to [n^y:], in a single step.

2.3. Postlexical palatalization

The most salient difference between all of the above processes and surface palatalization (of /t d n/ before /t^y d^y n^y j/) is that the latter does not involve coalescence: thus, e.g. *mit jelent* *[mit^y:elent] 'what does it mean', *védjegy* *[ve:d^y:ed^y] 'trade mark', *van joga* *[van^y:oga] 'he's got the right (to)'.

The simplest case of surface palatalization is where both target and trigger are noncontinuants (i.e., stops or nasals). This is where /t d n/ obligatorily (automatically, without exception) turn into [t^y d^y n^y] before /t^y d^y n^y/. Examples: *van gyufám* [n^{y}d^y] 'I've got matches', *két nyúl* [t^{y}n^y] 'two rabbits'.[10]

Examples like *hat tyúk* [hat^y:u:k] 'six hens', *mit gyártanak* [mid^y:a:rtanak] 'what do they produce' appear to contradict the claim we made above that surface palatalization does not involve coalescence. Actually, however, these are cases where the output of palatalization undergoes the rule of Long Consonant Formation (LCF: two adjacent identical short consonants merge into a single long one) either immediately (*hat tyúk*) or via voicing assimilation (*mit gyártanak*).

Thus, one branch of surface palatalization can be formulated as follows:

$$(18)\quad \begin{bmatrix} +\text{cor} \\ -\text{cont} \\ -\text{strid} \end{bmatrix} \longrightarrow [-\text{ant}] \;/\; \underline{\quad} \begin{bmatrix} -\text{ant} \\ +\text{cor} \\ -\text{strid} \\ -\text{cont} \end{bmatrix}$$

[10] Since the rule is postlexical, its application is not restricted to derived environments. Hence, cases like *satnya* 'stunted', *hangya* 'ant', *ponty* 'carp', etc. can be considered as derived by morpheme-internal applications of this rule (the relevant segments are underspecified for the feature [ant] and acquire the minus value via assimilation).

Before /j/, the surface palatalization of /t d n/ is optional: e.g. *két játék* [ke:t^{y}ja:te:k] ~ [ke:tja:te:k] 'two games', *szabad jönni* [sabad^{y}jön:i] ~ [sabadjön:i] 'you can come', *talán jobb* [talã:job:] ~ [tala:n^{y}job:] ~ [tala:njob:] 'perhaps better'.[11] This rule can be written as follows:

$$(19)\quad \begin{bmatrix} +\text{cor} \\ -\text{cont} \\ -\text{strid} \end{bmatrix} \longrightarrow [-\text{ant}] \;/\; \_\_ \begin{bmatrix} -\text{ant} \\ +\text{cor} \\ -\text{strid} \\ +\text{cont} \end{bmatrix}$$

These two rules, (18) and (19), can be collapsed as follows (the parenthesized [(−cont)] in the context is meant to suggest that, with increasing speed and/or casualness, that restriction is removed and the rule applies to /t d n/ before /j/ as well):

$$(20)\quad \begin{bmatrix} +\text{cor} \\ -\text{cont} \\ -\text{strid} \end{bmatrix} \longrightarrow [-\text{ant}] \;/\; \_\_ \begin{bmatrix} -\text{ant} \\ +\text{cor} \\ -\text{strid} \\ (-\text{cont}) \end{bmatrix}$$

Notice that (20) can also be used in lieu of (13) for word internal (lexical) palatalization (provided we ignore the asymmetry in optionality that we just noted and pretend, for simplicity's sake, that all postlexical applications are optional): in fact, there are a few marginal cases like *pattan* 'crack'—*pattantyú* 'cracker', *pillanat* 'moment'—*pillanatnyi* 'momentary' where the rule applies lexically in the context of palatals other than /j/. Thus, a unified rule of Palatalization will be proposed for both lexical and postlexical applications; it is, as expected, obligatory lexically and optional (i.e. rate/style-dependent) postlexically:

[11] In colloquial speech, a rule of *n*-vocalization (e.g. *színlap* [sĩ:lap] 'playbill', *tonhal* [tõ:hal] 'tuna', *kénsav* [kẽ:šav] 'sulphuric acid') usually bleeds surface palatalization of *nj* sequences; in guarded speech, on the other hand, neither rule is normally applied. There is, however, a narrow range of cases between [Ṽ:j] and [Vnj] realizations where forms of the [Vn^{y}j] type surface.

(21) Palatalization (lexical and postlexical)

$$\begin{bmatrix} +\text{cor} \\ -\text{cont} \\ -\text{strid} \end{bmatrix} \longrightarrow [-\text{ant}] \;/\; \underline{\quad} \begin{bmatrix} -\text{ant} \\ +\text{cor} \\ -\text{strid} \end{bmatrix}$$

(i.e. /t d n/ turn into [t^y d^y n^y] before /t^y d^y n^y j/)

Consider now the case of /l/. Across word boundary (including compound boundary) /l/ remains unaltered before /j/ in guarded speech; whereas in colloquial speech full coalescence takes place just like in lexical palatalization:

(22)	*hol jelent meg*	[holje-] ~	[hoj:e-]	'where did it appear'
	hiteljuttatás	[-elju-] ~	[-ej:u-]	'granting of credit'
	följut	[följut] ~	[föj:ut]	'reach the top'

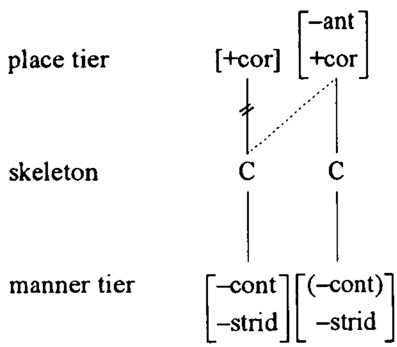
In casual speech, /l/ can be simply dropped (with or without compensatory lengthening of the preceding vowel: [fö:jut], [föjut]) before palatalization had the chance to apply. Before palatal noncontinuants, /l/ has the first and third options, but not the second (whether we interpret it as fusion or as palatalization to [j]):

(23)	*fél tyúk*	[fe:lt^{y}u:k]	~ [fe:t^{y}uk];	*[fe:t^y:uk],	*[fe:jt^{y}uk]	'half a hen'
	ökölnyi	[ökéln^{y}i]	~ [ökö:n^{y}i];	*[ökön^y:i],	*[ököjn^{y}i]	'fist-sized'
	kopoltyú	[kopolt^{y}u:]	~ [kopo:t^{y}u];	*[kopot^y:u],	*[kopojt^{y}u]	'gill'

All this suggests that the palatalization of *l* is a quite separate process and should not be collapsed with (21). Vago (1980, 40) assigns the feature value [−cont] to /l/ in order to be able to account for its behaviour in the context of /j/ with the general rule of Palatalization (roughly, our (21)). However, as we have just seen, /l/ does not palatalize before palatals other than /j/, either lexically or postlexically (this observation is due to Olsson 1992). Therefore, /l/ is taken to be [+cont], and has its own "palatalization" rule (actually, a rule of full assimilation to /j/). In fact, we have given that rule under (12) above; all we have to add at this point is that (12)—just like (21)—is obligatory lexically and optional/rate-dependent in its postlexical applications.

Finally, recall that we gave two alternative analyses for Lexical Palatalization in section **2.2.** The standard solution involved (13)—now recast as (21)—and (15), *j*-assimilation, applied in that order. The other account had a single rule of mutual assimilation, (17). Now if we opt for the second analysis, obviously a separate rule of Postlexical Palatalization is called for. This is essentially rule (20); it can be formalized in autosegmental terms as follows:

(24) Postlexical Palatalization (=(20):



The two analyses have different consequences with respect to the organization of phonology (into lexical and postlexical strata). These implications will be explored in the next section.

3. Implications for the stratal organization of phonology

As we observed in section **2.1**, *t*-palatalization/deletion must be restricted to imperative forms. One way we suggested this might be done (short of the diacritic marking of morphological information) is by assuming that the *j* of both personal and possessive suffixes is epenthetic and the relevant insertion rules are extrinsically ordered after *t*-pal/del (counterfeeding order). A more satisfactory solution (involving neither epenthesis nor extrinsic ordering) is to assume, in Lexical Phonology terms, that the imperative morpheme is one of the first-level (deeper) suffixes whereas personal suffixes belong to the second (more superficial) level (as is suggested by the order of their concatenation, too). If the rule of *t*-pal/del is restricted to Level 1, the /j/ of personal suffixes cannot trigger its application:

(25)	'let us see it'	'we wash it'	'we see it'
	[[[la:t]j]juk]	[[moš]juk]	[[la:t]juk]
Level 1	[la:t]j	–	–
(3) *t*-pal/del	š		
(15) *j*-assimilation	š		
Level 2	[la:šš]juk	[moš]juk	[la:t]juk
(21) Palatalization			t^y
(15) *j*-assimilation	š	š	t^y
Degemination	∅		
LCF	š:	š:	t^y:
OUTPUT	[la:š:uk]	[moš:uk]	[la:t^y:uk]

It is a characteristic feature of this solution that *j*-assimilation has to apply at both levels. It is impossible for it to start applying at Level 2 because of the Strict Cyclicity Principle (cf. Siptár 1988, 5 and references cited there). However, there is nothing wrong with a rule spanning several levels, as long as the levels concerned are contiguous. This condition is satisfied in the present case.

Moving on to lexical palatalization, recall that we suggested two solutions to account for cases like *látja* 'he sees it', *adja* 'he gives it', *bánja* 'he regrets it'. One involved a pair of rules, (21) Palatalization and (15) *j*-assimilation; the other posited a single rule for these cases: (17) Lexical Palatalization. Consider the classical generative account first. We noted that (21) Palatalization applies both lexically (in Level 2) and postlexically, whereas (15) *j*-assimilation is restricted to the lexical component (Levels 1 and 2). This asymmetry explains the fact that surface palatalization does not result in coalescence (except accidentally as in *hét tyúk* [t^y:] 'seven hens', *vad gyűlölet* [d^y:] 'fierce hatred', *vén nyárfa* [n^y:] 'old poplar'). Furthermore, (21) Palatalization cannot be excluded from Level 1, either: this is dictated by Strict Cyclicity again, given that the rule is also triggered by imperative *j* if the stem-final consonant is /d/ or /n/: *adjon* 'let him give', *kenjen* 'let him smear'. In accordance with Proper Inclusion Precedence, (3) *t*-pal/del is automatically ordered before (21) Palatalization; this explains why stem final /t/'s are never input to Palatalization in imperative forms. Consider the following derivations:

(26)	'let him give it'	'he gives it'	'momentary'
	[[[ad]j]ja]	[[ad]ja]	[[pillanat]n^{y}i]
LEXICAL PHONOLOGY			
Level 1	[ad]j	–	[pillanat]n^{y}i
(3) *t*-pal/del			
(21) Palatalization	d^y		t^y
(15) *j*-assimilation	d^y		
Level 2	[ad^{y}d^y]ja	[ad]ja	–
(21) Palatalization		d^y	
(15) *j*-assimilation	d^y	d^y	
POSTLEXICAL PHONOLOGY	[ad^{y}d^{y}d^{y}a]	[ad^{y}d^{y}a]	[pillanat^{y}n^{y}i]
(21) Palatalization			
Degemination	∅		
LCF	d^y:	d^y:	l:
OUTPUT	[ad^y:a]	[ad^y:a]	[pil:anat^{y}n^{y}i]

	'family law'	'stunted'	'old poplar'
	[[čala:d][jog]]	[šatn^{y}a]	[ve:n][[n^{y}a:r][fa]]
LEXICAL PHONOLOGY			
Level 1	–	–	–
Level 2	–	–	–
Level 3			
Compound Formation[12]	[čala:d][jog]	–	[n^{y}a:r][fa]
(21) Palatalization	d^y		
POSTLEXICAL PHONOLOGY	[čala:d^{y}jog]	[šatn^{y}a]	[ve:n][n^{y}a:rfa]
(21) Palatalization		t^y	n^y
LCF			n^y:
OUTPUT	[čala:d^{y}jog]	[šat^{y}n^{y}a]	[ve:n^y:a:rfa]

[12] The motivation for assigning Compound Formation to Level 3 is that *j*-assimilation does not apply across a compound boundary, e.g. *látjátok* [la:t^y:a:tok] 'you-pl. see it' (Level 2

We argued that /l/ should be excluded from the rule of Palatalization; rather, it should have its own rule of full assimilation to /j/. This move solves the dilemma of intermediate l^y and the problem that /l/ does not palatalize in any sense before /t^y d^y n^y/, at any level. The rule of *l*-palatalization is completely unrestricted as to which levels and components it applies in:

(27)	'let him live'	'devours it'	'comes up'	'where have you been'
	[[[e:l]j]en]	[[fal]ja]	[[fel][jön]]	[hol][[[ja:r]t]a:l]
LEXICAL PHONOLOGY				
Level 1	[e:l]j	–	–	[ja:r]t
(12a) *l*-palatalization	j			
Level 2	[e:jj]en	[fal]ja	–	[ja:rt]a:l
(12a) *l*-palatalization		j		
Level 3				
Compound Formation	–	–	[fel][jön]	–
(12a) *l*-palatalization			j	
POSTLEXICAL PHONOLOGY	[e:jjen]	[fajja]	[fejjön]	[hol][ja:rta:l]
(12a) *l*-palatalization				j
LCF	j:	j:	j:	j:
OUTPUT	[e:j:en]	[faj:a]	[fej:ön]	[hoj:a:rta:l]

The autosegmental equivalent of the above derivations runs as follows (transcription symbols stand for root nodes with all the material they dominate):

morphology: both Palatalization and *j*-assimilation apply) vs. *átjárok* *[a:t^y:a:rok] 'I (often) go through' (Level 3 morphology (compounding): Palatalization optionally applies ([tj] ~ [t^yj]) but *j*-assimilation does not). Palatalization has to appear on Level 3, too; the only way for it to appear on Levels 1–2 *and* postlexically (and observe the condition of level contiguity referred to above) is to appear on Level 3 as well. As far as the data go, however, it would be possible for Palatalization to be absent from Level 3, given that Strict Cyclicity is inoperative in the postlexical component, thus palatalization across compound boundary could be covered by postlexical palatalization.

(28)	'let him live'	'devours it'	'comes up'	'where have you been'
	[[[VVC]C]VC] V | | || e l j en	[[CVC]CV] ||| || fal ja	[[CVC][CVC]] ||| ||| fel jön	[CVC][[[CVVC]C]VVC] ||| |V| |V| hol j a r t a l
LEXICAL PHONOLOGY				
Level 1	[VVC]C V | | e l j	–	–	[CVVC]C |V| | j a r t
(12b) *l*-palat.	VVC C V V e j	–	–	–
Level 2	[VVCC]VC V V || e j en	[CVC]CV ||| || fal ja	–	[CVVCC]VVC |V ||V| j a rt a l
(12b) *l*-palat.	–	CVC CV ||V| fa j a	–	–
Level 3 Compound Formation	–	–	[CVC][CVC] ||| ||| fel jön	–
(12b) *l*-palat.	–	–	CVC CVC ||V|| fe j ön	–
POSTLEXICAL PHONOLOGY	[VVCCVC] V V || e j en	[CVCCV] ||V| fa j a	[CVCCVC] ||V|| fe j ön	[CVC][CVVCCVVC] ||| |V||V| hol j a rt a l
(12b) *l*-palat.	–	–	–	CVC CVVCCVVC ||V V||V| ho j a rt a l
OUTPUT	[e:j:en]	[faj:a]	[fej:ön]	[hoj:a:rta:l]

Notice that LCF is not involved in these derivations: unlike (12a), the autosegmental rule (12b) creates linked structures (i.e., geminate [j:]'s) directly. This does not mean, however, that an autosegmental phonology of Hungarian can do without LCF. Cases where two identical short consonants become adjacent by simple juxtaposition (i.e., without any kind of assimilation) require LCF in any framework (e.g. [[hat]to:l] ⟶ [hat:o:l] *hattól* 'from six', [hol][les:] ⟶ [hol:es:] *hol lesz* 'where will it be').[13]

The role of Degemination is also different in the two frameworks. In classical-generative derivations (cf. (25), (26)) we assumed, for simplicity's sake, that Degemination applied (only) in the postlexical component. This resulted in intermediate forms like *la:šššuk* 'let us see it' (the output of Level 2 in (25)) or *$ad^yd^yd^ya$* 'let him give it' (the output of Level 2 in (26)), or even *hajjjjuk* 'let us hear it' (<[[[hall]j]juk]). This is rather counterintuitive, but in the classical framework it works, and is the simplest solution.[14] In the autosegmental framework, however, there is an important formal reason for Degemination to apply as the first rule of both Levels 1 and 2 (in addition to its postlexical application). Take the example of *halljuk* 'let us hear it'. The underlying form is [[[CVCC]C]CVC]; the Level 1 input is [CVCC]C. Now,

```
[[[CVCC]C]CVC]        [CVCC]C
   | | V | | | |         | | V |
   h a l j j u k         h a l j
```

this form should undergo (12b) *l*-palatalization but it cannot, since it does not match the structural description of the rule:

```
C C
| |
l j
```

(In order for an autosegmental rule to apply, its structural description has to be matched in an **exhaustive** manner; thus, it is not only the case that a rule requiring

```
C C C
 V  |
 l  j
```

[13] On the other hand, LCF is an automatic consequence of the universal Obligatory Contour Principle (OCP, cf. McCarthy 1986, 208) in this framework. Hence, it is "free of charge": it does not make the grammar more complicated. The principle can be formulated as follows:

OCP: On the melodic (=non-skeletal) tiers, adjacent identical elements are prohibited.

In some languages the OCP triggers LCF as in Hungarian; in others, it acts as a filter and removes offending structures (or does not let them come into being). In the latter type of languages, long consonants and/or long vowels are disallowed (at some level). For instance, in English, geminates are not permitted lexically (at Levels 1 and 2) but do occur across compound boundary (and postlexically) via LCF; cf. the familiar cases like *unknown*, *misspell*.

[14] That simplicity may, however, turn out to be spurious in that a lexical (Level 2) and a postlexical type of Degemination will have to be distinguished in any framework (cf. Nádasdy 1989 for details). This issue will not be pursued here.

cannot apply to a form involving C C; the same is true the other way round.)
| |
l j

But if we assume Degemination to apply before (12b), we get a proper derivation [CVCC]C —(DEG)⟶ [CVC]C —(12b)⟶ [CVC]C. But this is not
| | V | | | | | | | V
h a l j h a l j h a j

the whole story. In a form like *adjuk* 'let us give it', the input to Level 2 is [VCC]CVC. Again, Degemination has to turn this into [VC]CVC before (17)
| V | | | | | | | |
a d^y j u k a d^y j u k

Lexical Palatalization can apply (giving [VC]CVC). Hence, it is not only at
| V | |
a d^y u k

the beginning of Level 1, but also at that of Level 2, that Degemination has to apply. Let us call this type of degemination (i.e., the one that is lexical, obligatory, and partly motivated by the framework itself) 'Degemination I'. In addition, we need 'Degemination II', a postlexical rule that is optional/rate-dependent, and subject to the restrictions and qualifications discussed in Nádasdy (1989). The two rules can be informally written as in (29a,b); the formulation (29c) is the classical-generative rule of Degemination, added here for comparison.

(29) Degemination

(a) C C ⟶ C / ___ C (Levels 1 and 2)
V |

$$\text{(b)}\quad \underset{\bigvee}{\mathrm{C\,C}} \longrightarrow \underset{|}{\mathrm{C}} \;//\; \_\_\_ \begin{Bmatrix} \begin{matrix} \mathrm{C} \\ / \\ [-\mathrm{son}] \end{matrix} & \text{(i)} \\ \begin{matrix} \mathrm{C} \\ / \\ [+\mathrm{nas}] \end{matrix} & \text{(ii)} \end{Bmatrix} \qquad \text{(Postlexical)}$$

(c) $C_i \longrightarrow \emptyset \;//\; C_i \_\_\_ C$ (Postlexical, "optional")

Note that (29c) is ordered **before** LCF, whereas (29b) is ordered **after** LCF. C_i in (29c) stands for an arbitrary consonant whose feature values exactly match those of the other consonant marked C_i; // means 'in the environment specified and its reverse' (mirror image). "Optional" in (29c) loosely refers to a state of affairs that the rule is obligatory in some circumstances and style/rate-dependent in others; cf. Nádasdy (1989) for details and examples. The autosegmental version spells out the pattern of optionality more in detail: (29a), being a lexical rule, is *a priori* obligatory (e.g. *hallva* 'hearing it'); (29b(i)) is obligatory in colloquial speech (at normal speed), whereas (29b(ii)) is optional there (becoming obligatory in casual/fast speech only). Cf. *koszttól* 'from food', *kisstílű* 'petty' (29b(i)): formal [st:], [š:t], colloquial/casual [st], [št]; *tankként* 'like a tank', *kész sznob* 'a perfect snob' (29b(ii)): formal [ŋk:], [s:n], colloquial [ŋk(:)], [s(:)n], casual [ŋk], [sn]; cf. *talppont* 'nadir', *széppróza* 'prose fiction': [lp:], [p:r]. In fast casual speech, (29b) will generalize into $\underset{\vee}{\mathrm{C\ C}} \longrightarrow \underset{|}{\mathrm{C}}$ // __ C (*talppont* [lp], *széppróza* [pr]), and eventually into $\underset{\vee}{\mathrm{C\ C}} \longrightarrow \underset{|}{\mathrm{C}}$ (i.e., any geminate may shorten, even intervocalically).

In the autosegmental account, then, there is no general all-purpose Palatalization rule (like (21) in the classical version). Only one rule spans all lexical levels, as well as the postlexical component: *l*-palatalization (12b). The other relevant rules are either purely lexical rules, restricted to Level 1 ((8) (10), (11)) or to Levels 1 and 2 ((9), (17), (29a)), or else they are purely postlexical rules ((24), (29b), LCF). Consider a few sample derivations that illustrate the rules discussed in **2.1** (cf. (30)), **2.2** (cf. (31)), and **2.3** (cf. (32)), respectively.

(30)	'let it please'	'let him divide it'	'let him see it'	'let him heat it'
	[[[CVCC]C]VVC] t e t s j e k	[[[VCC]C]CV] o s t j j a	[[[CVVC]C]CV] l a t j j a	[[[CVCC]C]CV] f ü t j j e
LEXICAL PHONOLOGY				
Level 1	[CVCC]C t e t s j	[VCC]C o s t j	[CVVC]C l a t j	[CVCC]C f ü t j
(8) *t/s* → ∅	CVCC*C t e t j	VCC*C o s j	–	–
(10) *t*-pal[č]	–	–	–	CVCC C f ü č
(11) *t*-pal[š]	CVCC*C t e š	–	CVVCC l a š	–
(9) *j*-assim.	–	VCC*C o s	–	–
(17) Lex. Pal.	–	–	–	–
Level 2	[CVCC]VVC t e š e k	[VCC]CV o s j a	[CVVCC]CV l a š j a	[CVCCC]CV f ü č j e
(29a) Degemination	–	VCCV o s j a	CVV C CV l a š j a	CVC C CV f ü č j e
(9) *j*-assim.	–	VCCV o s a	CVV C CV l a š a	CVC C CV f ü č e
(17) Lex. Pal.	–	–	–	–
POSTLEXICAL PHONOLOGY	[CVCCVVC] t e š e k	[VCCV] o s a	[CVVCCV] l a š a	[CVCCCV] f ü č e
(24) Postlex. Pal.	–	–	–	–
OUTPUT	[teš:e:k]	[os:a]	[la:š:a]	[fü:č:e]

(31)

<table>
<tr><td></td><td>'let him give it'</td><td>'gives it'</td><td>'sees it'</td><td>'throws it'</td></tr>
<tr><td></td><td>[[[VC]C]CV]
| | | | |
ad j j a</td><td>[[VC]CV]
| | | |
ad j a</td><td>[[CVVC]CV]
| V | | |
l a t j a</td><td>[[CVVC]CV]
| V | | |
h a n^y j a</td></tr>
<tr><td>LEXICAL PHONOLOGY</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Level 1</td><td>[VC]C
| | |
ad j</td><td>–</td><td>–</td><td>–</td></tr>
<tr><td>(17) Lex. Pal.</td><td>VC C
| V
a d^y</td><td></td><td></td><td></td></tr>
<tr><td>Level 2</td><td>[VCC]CV
| V | |
a d^y j a</td><td>[VC]CV
| | | |
ad j a</td><td>[CVVC]CV
| V | | |
l a t j a</td><td>[CVVC]CV
| V | | |
h a n^y j a</td></tr>
<tr><td>(291) Degemination</td><td>V C CV
| | | |
a d^y j a</td><td>–</td><td>–</td><td>–</td></tr>
<tr><td>(17) Lex. Pal.</td><td>V C CV
| \/ |
a d^y a</td><td>VC CV
| \/ |
a d^y a</td><td>CVVCCV
| V V |
l a t^y a</td><td>CVVCCV
| V V |
h a n^y a</td></tr>
<tr><td>POSTLEXICAL PHONOLOGY</td><td>[VCCV]
| \/ |
a d^y a</td><td>[VCCV]
| \/ |
a d^y a</td><td>[CVVCCV]
| V V |
l a t^y a</td><td>[CVVCCV]
| V V |
h a n^y a</td></tr>
<tr><td>(24) Postlex. Pal.</td><td>–</td><td>–</td><td>–</td><td>–</td></tr>
<tr><td>OUTPUT</td><td>[ad^y:a]</td><td>[ad^y:a]</td><td>[la:t^y:a]</td><td>[ha:n^y:a]</td></tr>
</table>

4. Conclusion

In this paper, we have given two alternative accounts of palatalization rules in Hungarian. The classical-generative account involved the rules listed in (33), overleaf.

With respect to these rules, the following rule interactions (ordering relationships) obtain: a bleeding/counterbleeding (i.e., mutual bleeding) relationship between (3) and (21) at Level 1; the attested ordering is determined by Proper Inclusion Precedence (PIP). (21) feeds (15), which in turn feeds (29c) and LCF. (12a) is not crucially ordered with respect to any of the lexical rules; but it

<table>
<tr><td>(32)</td><td>‘we’ve got matches’</td><td>‘six hens’</td><td>‘he’s got the right (to)’</td><td>‘trade mark’</td></tr>
<tr><td></td><td>[CVC][C VCV]
| | | | | | |
van d^yufa</td><td>[CVC][C VVC]
| | | | V |
hat t^y uk</td><td>[CVC][[CVC]V]
| | | | | | |
van joga</td><td>[[CVVC][CVC]]
| V | | | |
ved jed^y</td></tr>
<tr><td>LEXICAL PHONOLOGY
Level 1</td><td>–</td><td>–</td><td>–</td><td>–</td></tr>
<tr><td>Level 2</td><td>–</td><td>–</td><td>[CVC]V
| | | |
joga</td><td>–</td></tr>
<tr><td>Level 3
Compound Formation</td><td>–</td><td>–</td><td>–</td><td>[[CVVC][CVC]
| V | | | |
ved jed^y</td></tr>
<tr><td>POSTLEXICAL PHONOLOGY</td><td>[CVC][C VCV]
| | | | | | |
van d^yufa</td><td>[CVC][C VVC]
| | | | V |
hat t^y uk</td><td>[CVC][CVCV]
| | | | | | |
van joga</td><td>[CVVCCVC]
| V | | | |
vedjed^y</td></tr>
<tr><td>n-vocalization</td><td>–</td><td>–</td><td>CVC CVCV
| V | | | |
vã joga</td><td>–</td></tr>
<tr><td>(24) Postlex. Pal.</td><td>CVC C VCV
| | | | | | |
van^y d^yufa</td><td>CVC C VVC
| | \/ V |
ha t^y uk</td><td>–</td><td>CVVC CVC
| V | | | |
ved^yjed^y</td></tr>
<tr><td>OUTPUT</td><td>[va$n^y$$d^y$ufa]</td><td>[ha$t^y$:u:k]</td><td>[vã:joga]</td><td>[ve:d^yjed^y]</td></tr>
</table>

feeds LCF. Finally, (29c) and LCF are in a mutual bleeding relationship; (29c) precedes LCF due to PIP.

(33)	*t*-pal/del (3)	Level 1
	Palatalization (21)	Levels 1–3 + Postlexical
	j-assimilation (15)	Levels 1–2
	l-palatalization (12a)	Levels 1–3 + Postlexical
	Degemination (29c)	Postlexical
	Long Consonant Formation	Postlexical

The autosegmental rules we have posited can be summarized as follows:

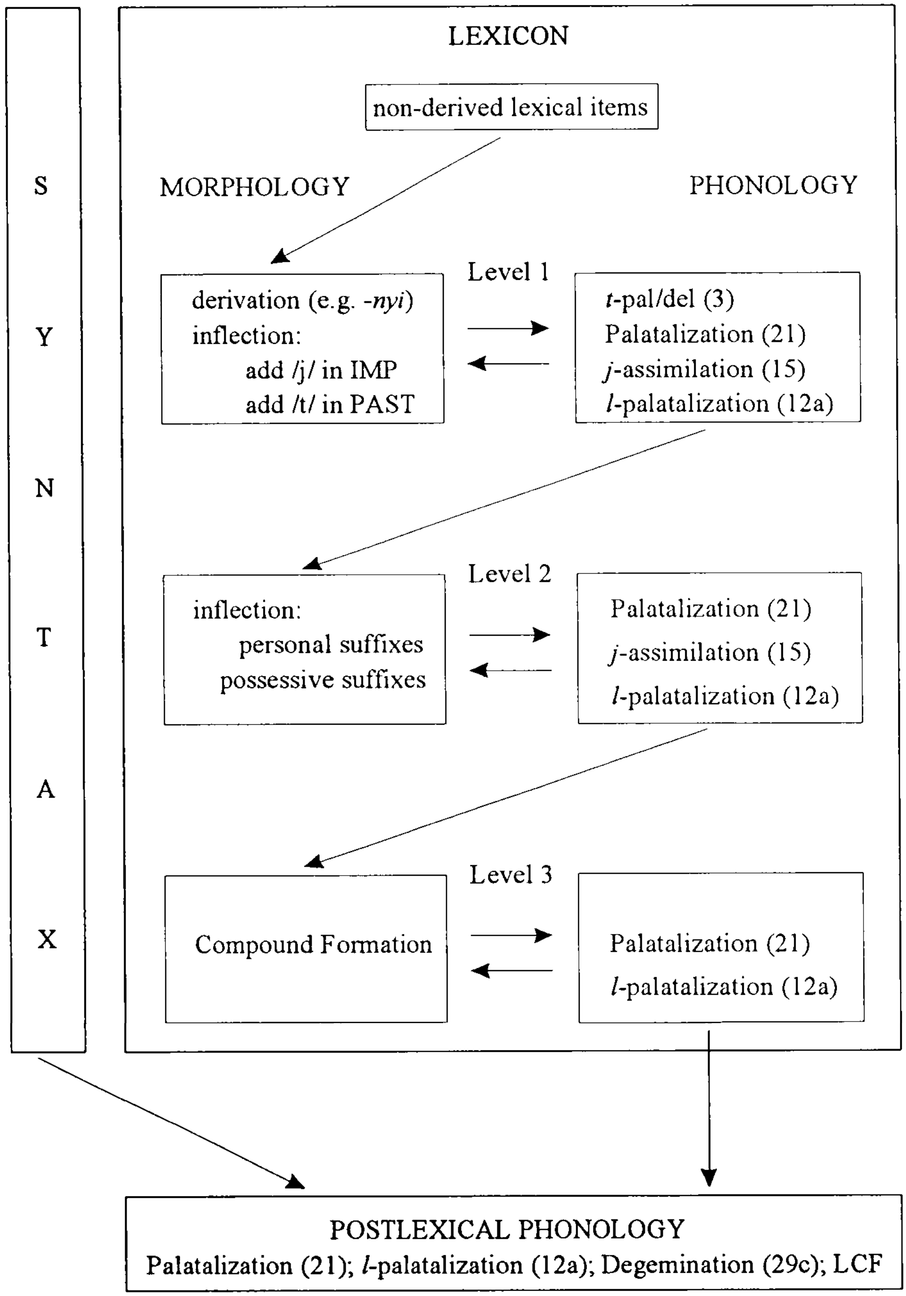
(34)	Degemination I (29a)	Levels 1–2	
	t/s-deletion (8)	Level 1	
	t-pal[č] (10)	Level 1	
	t-pal[š] (11)	Level 1	
	j-assimilation (9)	Levels 1–2	
	Lexical Palatalization (17)	Levels 1–2	
	l-palatalization (12b)	Levels 1–3 +	Postlexical
	Postlexical Palatalization (24)		Postlexical
	Long Consonant Formation		Postlexical
	Degemination II (29b)		Postlexical

There is a counterbleeding relationship between (8) and (10), between (10) and (11), as well as between (8) and (11); in addition, rule (8) feeds (9) and (11), (29a) feeds (9) and (17), and LCF feeds (29b). Each member of the set (8), (10), (11) is in a mutual bleeding relationship with (17); precedence is again as predicted by PIP. Finally, (12b) is not crucially ordered with respect to any other rule.

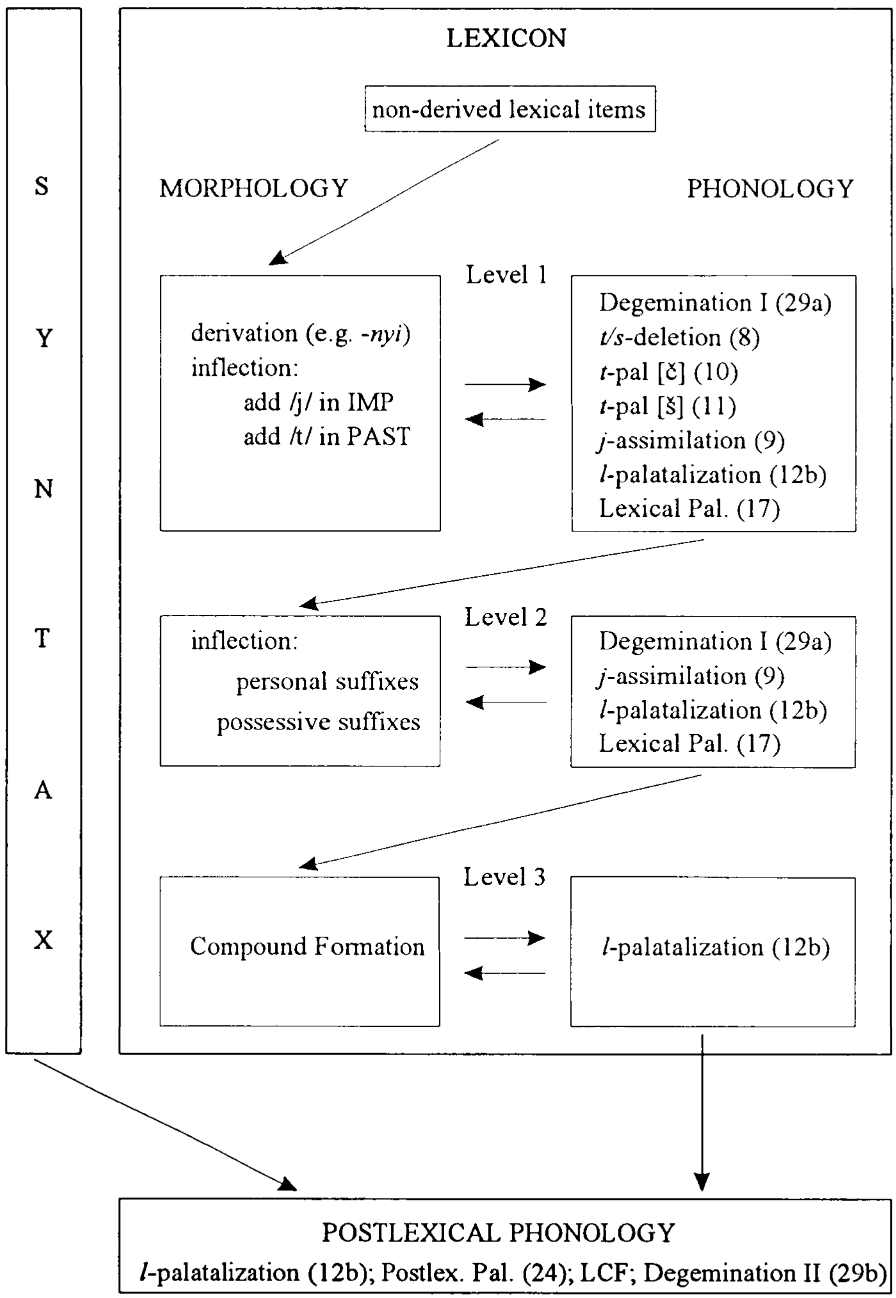
The classical–generative solution involves a total of six rules; the autosegmental solution comprises ten rules altogether. However, the difference is apparent rather than real: if we uncollapse (3) into three subrules and (15) into two, as well as introduce a rule of Lexical Degemination corresponding to (29a) (cf. footnote 14), the number of rules will be exactly the same in both frameworks. On the other hand, some of the autosegmental rules account for more than their classical–generative counterparts: (8) covers *s*-deletion as well as *t*-deletion; (24) Postlexical Palatalization captures the asymmetry in optionality as noted in (18)-(20) whereas the postlexical application of (21) Palatalization has to ignore it; similarly, with respect to Degemination, (29a-b) represent the pattern of optionality more exactly than does (29c); and the individual derivations are invariably shorter in the autosegmental case because coalescence rules typically do the job of two or three classical–generative rules each. For example, *falja* 'he devours it' involves just one rule in (28) but two in (27); *adja* 'he gives it' involves one rule in (31) but three in (26); *lássa* 'let him see it' involves three rules in (30) but five in (25), and so forth.

The two analyses are diagrammatically summarized in (35) and in (36), respectively.

(35) The classical–generative analysis:



(36) The autosegmental analysis:



References

Abondolo, D. 1988. Hungarian Inflectional Morphology. Akadémiai Kiadó, Budapest.

Clements, G.N. 1985. The geometry of phonological features. In: Phonology Yearbook 2: 225–52.

Deme, L. 1961. Hangtan [Phonetics and phonology]. In: Tompa, J. (ed.): A mai magyar nyelv rendszere [The system of present-day Hungarian]. Vol. I, 57–119. Akadémiai Kiadó, Budapest.

Jakab, L. 1967. A *t* végű igék felszólító módjához [The imperative of *t*-final verbs]. In: Magyar Nyelv 63: 194–6.

Kiefer, F. 1985. The possessive in Hungarian: A problem for Natural Morphology. In: Acta Linguistica Hungarica 35: 85–116.

Kontra, M. 1992. Fonológiai általánosítás és szociolingvisztika realitás [Phonological generalization and socioliguistic reality]. In: Kontra, M. (ed.): Társadalmi és területi változatok a magyar nyelvben [Social and geographical varieties in Hungarian]. Linguistica, Series A, Studia et Dissertationes 9, 87–95. Institutum Linguisticum Academiae Scientiarum Hungaricae, Budapest.

McCarthy, J. 1986. OCP effects: gemination and antigemination. In: Linguistic Inquiry 17: 207–63.

Nádasdy, Á. 1989. The exact domain of consonant degemination in Hungarian. In: Szende, T. (ed.): Proceedings of the Speech Research '89 International Conference, 104–7. Linguistics Institute of the Hungarian Academy of Sciences, Budapest.

Olsson, M. 1992. Hungarian Phonology and Morphology. University Press, Lund.

Siptár, P. 1988. Cyclicity and palatalization. In: Magyar Fonetikai Füzetek/Hungarian Papers in Phonetics 19: 5–13.

Siptár, P. 1993. Hungarian consonants: rules and representations. In: Kövecses, Z. (ed.): Voices of Friendship. Essays in Honor of László T. András, 27–50. Eötvös Loránd University, Budapest.

Vago, R.M. 1980. The Sound Pattern of Hungarian. Georgetown University Press, Washington.

Vago, R.M. 1987. On the representation of length. In: Dressler, W. *et al.* (eds): Phonologica 1984, 319–24. Cambridge University Press, Cambridge.

Vago, R.M. 1989. Compensatory lengthening in Hungarian: Implications for mora theory. Ms. Queens College, City University New York.

Vago, R.M. 1991. Elméleti szempontok a *t* végű igék elemzésében [Theoretical aspects of the analysis of *t*-final verbs]. In: Kiss, J.–Szűts, L. (eds): Tanulmányok a magyar nyelvtudomány történetének témaköréből [Papers on the history of Hungarian linguistics], 682–90. Akadémiai Kiadó, Budapest.

Address of the author: Péter Siptár
Research Institute for Linguistics
Hungarian Academy of Sciences
Színház u. 5–7.
1250 Budapest, P.O. Box 19.
Hungary

Acta Linguistica Hungarica, Vol. 42 (1-2), pp. 33-50 (1994)

THE EFFECT OF SYNTAGMATIC CONTEXT ON LEXICAL ACCESS*

MÁRIA GÓSY

Introduction

The process by which a speaker produces an utterance or a listener understands one involves much more than the generation or the auditory processing of a signal (Stevens 1978). When understanding speech, the listener uses his knowledge of the words of his language, which includes knowledge of what sequence of phonemes represents each word, and knowledge of the semantic and syntactic properties of those words. The questions are how information in the speech wave makes contact with the lexicon, and what kind of information about words is used in this proces.

Several studies have dealt with the problem of word recognition and lexical access and, as a consequence, numerous models have been proposed to describe the process (Klatt's review: 1989; Altmann 1991; etc.). Researchers have tended to focus on the recognition of isolated words spoken in their citation form and, simplifying the problem in this way, they have ignored important facts like the role of prosody, phonological processes, semantic and syntactic factors, the segmentation problem (cf. Cutler–Butterfield 1990). Segmentation, however, seems to be far more curious in certain languages where—due to their agglutinating character—segmentation should be performed at two levels: within word-like sequences and between them. In other words, if we take a continuously uttered Hungarian sentence-like sequence: *ezérthatalmasok*, we will have two segmentation problems as a consequence of the continuous acoustic wave to be perceived.

	Segmentation possibility	Semantic outcome
(i)	Ezért/hatalmasok.	For this reason/they have power.
(ii)	Ezért/hat/alma/sok.	For this reason/six/apples/are a lot.
(iii)	Ezért/hatalma/sok.	For this reason/his power/is excessive.
(iv)	Ezért/hatalmas/ok.	For this reason/it is a mighty/cause.

* This research was supported by OTKA [Project T 013863].

1216–8076/94/$ 4.00

One type of segmentation concerning longer chunks of the sequence can be called 'external' segmentation while the other type concerning the segmentation possibilities within the word can be called 'internal' segmentation. Naturally, prosodic information might be of great help for the distinction of 'external' and 'internal' segmentations, however, further experiments are necessary to detect which actual elements of the acoustic structure are responsible for prosodic information and how they are used in perceptual judgements.

On the word recognition process

It has generally been claimed that word recognition is based on acoustic structure on the one hand and on word frequency on the other (cf. Solomon–Postman 1952; Broadbent 1967; Nooteboom–Truin 1980; Elliott–Clifton–Servi 1983; Kintsch–Mross 1985; etc.). The context-dependency of word recognition has also been highlighted (cf. Miller–Heise–Lichten 1951; Fry–Denes 1952; Ladefoged 1967; Tanenhaus–Lucas 1987; Zwitserlood 1989; etc.). Models of word recognition during the last decade became much more concerned with the role of the statistical properties of the lexicon in the ease of access of a word (cf. Samuel 1986; Johnson-Laird 1987; Forster 1989).

What do we definitely know about word recognition? The assumptions are:

- Words are recognized through the interaction of sound and knowledge (e.g. Cole–Jakimik 1980).

- Though speech is generally processed sequentially, word by word (in production); word recognition in fluent speech is not necessarily processed sequentially, word by word.

- Syntactic and semantic contrasts are used to restrict the number of word candidates.

- The prosodic structure of speech is often highly involved in the recognition process (cf. Slowiaczek 1990); however several apparently negative indications have been found (using English data) against the role of stress in the word recognition process (cf. Cutler 1986).

- There are some problems that tend to be emphasized recently like the mental lexicon (e.g. Schreuder–D'Arcais 1989).

Terminology used in the literature is often confusing and contradictory because two terms are used for almost the same process (cf. Frauenfelder–Komisarjevsky-Tyler 1987). One of these is "word recognition" which refers to processes responsible for generating a pattern from the acoustic–phonetic information in the speech waveform and matching this pattern to patterns previously stored in memory (i.e. for words). The other one is "lexical access" which refers to processes that are responsible for contacting the appropriate lexical information in memory once a pattern match has been accomplished. However, it is very difficult to separate the processes of word recognition and lexical access from their products. There seems to be an agreement in the literature concerning the two stages of lexical processing: (i) At the first stage the contact between the sensory input and the candidate lexical representations is established. (ii) At the second stage certain postlexical processes apply which involve the selection, elaboration, and integration with the context of the information made available at the first stage. In other words: the result of the operations at the first stage is a possible form of an expected (virtual) word, and at the second stage the output is an actual word having acoustic, phonetic, syntactic, semantic, and prosodic patterns appropriate to the context.

There are numerous models concerning word identification and lexical access and also a lot of hybrid models developed on the basis of some current hypotheses. There are two trends in the models: (i) one is the modularity hypothesis and (ii) the second type is the interactive model (cf. Frauenfelder–Komisarjevsky-Tyler, eds. 1987). Recent'y, accepting the existence of the two types of models, a novel distinction can be found in the literature. According to this classification, one type of models is called activation while the other type is called lexical search models (cf. Forster 1989). The most prominent word recognition models are frequently distinguished by their position on a continuum ranging from fully interactive to completely autonomous (e.g. Morton 1969; Forster 1979; Seidenberg 1985; McClelland–Elman 1986; Marslen-Wilson 1989).

Independently of the different labellings, the philosophy of distinctions between the two trends of models seems to be the same: in one type (modularity or activation hypotheses), modules communicate only at the interfaces between levels; the output of one module may serve as the input to a second module. However, the internal operations of a module are blind to what is going on in other modules; the processes are assumed to take place parallel in time. In these models every word (or logogen or node) is a comparator for a candidate list with a single member and is represented by all necessary patterns and strategies for continuous comparisons between the acoustic input and the mental lexicon. The centres of operations are the lexical units (i.e. the words) themselves. The connectionistic approach to word recognition has basically grown out of these theories.

In contrast, interactive (or lexical search) models predict that context can influence what information is accessed during lexical processing. Operations at various levels of the mechanism interact with one another, interaction provides the possibility of permanent correction, and these continuous processes take place real-time (though parallel operations are not excluded). These models can be characterised by serial comparisons whose results are gathered in a central comparator. The comparator in question is interrelated both with the acoustic input and the lexical units.

According to the first type of hypothesis, words compare themselves to the input acoustic signal while in the second type of hypothesis this task is given to the central comparator which compares the acoustic input to a group of words in the mental lexicon. There is a very common problem in both models: how to differentiate among acoustically very similar sequences. In lexical search models this is meant to be solved by emphasizing the factor of frequency while in activation models a so-called activation weight is supposed to work by means of a specific threshold value.

These models provide a mechanism for higher-level knowledge to guide lower-level processing whereas modular or parallel models basically deny the existence of such feedback.

In all models dealing with word recognition attempts have been made to investigate the function of various factors, like semantic and syntactic factors (Feldman 1991), phonological rules and processes (Frauenfelder *et al.* 1990), memory, frequency of occurrence, the familiarity of words, the role of predictions and expectations, the role of the prosodic level and variability as a function of age. The problems that have remained open are the following: how do these factors actually function during recognition and how are language specific patterns integrated into the process?

Word recognition is based on acoustic–phonetic information while in context both semantic and syntactic components are used to deactivate some possible word candidates (Klatt 1980; Seidenberg *et al.* 1982; Pisoni–Luce 1987). Context can serve to narrow down the available range of choice so that only a small portion of the acoustic waveform needs to be identified for word recognition to take place successfully. However, it is very difficult to say how context-dependency works, what the most important factors are in this processing, what the similarities and differeneces are in word recognition and lexical access between isolated words and words in context. If we look at the various models, it can be seen that highly intricate ways of interactions are postulated among the levels of processing. According to the findings of Tanenhaus–Lucas (1987), there is strong evidence for lexical context effects, mixed evidence for semantic effects and little evidence for syntactic effects. The present author wants to express her doubts concerning these claims since: (i) lexical effects normally operate prior to both semantic and syntactic effects, (ii) semantic and syntactic effects operate differently depending on the size of the context, (iii) effects of the syntactic factors depend on the language structure in question, and (iv) syntactic effects in agglutinating languages should have a twofold function which seems to be eliminated in the description of models that are mainly based on English material and experiments with English-speaking subjects.

The process of recognizing a spoken word begins when sensory input makes initial contact with the lexicon. In this phase the listener perceives the acoustic structure of the speech sample as input and generates the representation(s) which contact the internally stored representations associated with each lexical entry. The listeners generally recognize the words both in isolation and in context before having heard them completely. According to Marslen-Wilson's cohort theory (1989)—which is based on this hypothesis—, acoustic–phonetic information at the beginning of an input word activates all words in one's memory that share this word-initial information, known as "the cohort". Lexical access mechanism in cohort theory works from left to right on a narrow phonetic or featural representation as opposed to a phonemic representation of the input. At this stage of the model, word recognition is a completely bottom-up process. At the second stage, a particular word will be recognized at the 'critical recognition point' where the word is uniquely distinguished from any other word in the language beginning with the same initial sound sequence; which means that the cohort reduces to one word, a lexical decision is rendered, even if the acoustic end of the word has not been heard. Context plays an important role in this second stage in reducing the size of the set of activated words. The recognition of a word can have an impact upon the processing of another word which bears some relationship to the first one. This relationship can be either structural, i.e. syntactic, or non-structural, i.e. semantic. This model practically denies the interaction of (possible) top-down operations with the bottom-up accessing mechanism (Marslen-Wilson 1989).

Our hypothesis concerning the word recognition process from a non-connectionist approach contains the following assumptions that partly agree with, and partly contradict, those proposed in the literature.

We think that the operations of the so-called lower levels of the word recognition process are much more autonomous than those of the higher levels. An actual model should be highly interactive, operating flexibly among the hypothesized levels. Namely, both possible accessing mechanisms: top-down and bottom-up should be at work, with the priority of bottom-up operations. Invariance should be possible at all levels while its appearance is based on the actual perception/comprehension task (cf. Gósy 1992). The input is a bottom-up analysis while feedbacks could take place between the higher and lower levels, resulting in continuous alterations, modifications and changes in the decisions at each level. Our claim is that, given a lexical hypothesis, the interactive operations (with the possibility of parallelism) and the continuous verification mechanism, by obtaining newer and newer "perceptual" information, will result in an unambiguous final lexical decision (which is supposed to be right). The only factor restricting this process is the temporal limit for all operations.

A series of experiments (with Hungarian words and listeners having Hungarian as their first language) was conducted in order to answer the question of primacy in word recognition for either acoustic patterns or word frequency (Gósy 1987; 1992). The results have suggested that the recognition of words is based primarily on the processing of their acoustic structure; however, there is no doubt that other factors like semantic associations, frequency, or (to a lesser degree) the subject's personality might also influence the final decision. These investigations have given evidence that the language-specific phonotactic rules are encoded into the mental lexicon and also play a great role in word recognition. In one of the experiments meaningless sound sequences masked by white noise were to be identified—according to the instructions—as real Hungarian words. The meaningless sound sequences with various acoustic structure either obeyed or contradicted the Hungarian language-specific phonotactic rules. The listeners' reaction time, i.e. the amount of time that elapsed between the target sequence and the subjects' repetition, was measured in milliseconds. The results have significantly shown that phonotactic patterning has a very important role in recognition (particularly vowel harmony that is fairly pervasive in Hungarian). The increase of reaction time for meaningless items is parallel to the increase of divergence of these items from Hungarian phonotactic patterns.

Meaningful Hungarian words are characteristically identified in a way that a decision is taken on the basis of initial portion(s) of the acoustic input, a decision that usually turns out to be correct. This strategy has been termed 'progressive identification' (cf. Gósy 1992) where the whole acoustic pattern is not processed since correct identification is made possible by the initial portion(s). However, if the listener has some difficulty with the operations in the process somewhere the acoustic structure will have to be analyzed in full so that meaning assignment can be carried out in possession of as much information as possible. In this case very frequently the final part of the sound sequences is stored correctly and the final decision on the word is made on the basis of this part. This procedure is called 'regressive identification' characteristic mainly of the recognition of meaningless sound sequences and may also appear in the case of real words. For example, the average reaction time for the recognition of the original word *verem* 'pit' was 309 ms while the average reaction time for the same word recognized regressively as *merev* 'stiff' was 426 ms. In all cases, reaction times in such 'regressive identification' were significantly longer than those in 'progressive identification'. On the basis of these findings it seems that word recognition does not operate necessarily in a strictly left-to-right fashion, though this appears as unquestionable in all models. It is likely that that assumption goes back to earlier results of word recognition experiments when research focused on written words. And, this left-to-right mood of operation could also obtain some support from phonology as well as from the fact that words are formed, in articulation, by a string of speech sounds constructed progressively in time. However, this does not mean that word recognition should equally be processed progressively in time. We are aware that our contention violates unquestioned claims about lexical access; however, it might be supposed that the nature of an agglutinating language makes the listeners' decoding system flexible enough for bidirectional recognition. This could also be explained by one of the universal rules of language acquisition: pay attention on the ends of the words (Slobin 1970) which is very important also for adults (particularly for those who speak an agglutinating language).

There are three types of evidence for the reversed direction of recognition. (i) One is the ability of speaking fluently backward, which has long been known to exist (generally as a secret language in childhood). Analysis of backward speech showed evidence of sensitivity to bidirectional phonemic representations of the utterances (Cowan *et al.* 1982). (ii) There is an increasing number of children producing metatheses in their speech around the age of 6, although they speak and perceive normally. Most of them have no actual dominance in using either the right or the left hand. (iii) It is well known that young children acquiring their mother tongue imitate either the beginning or the final part of the word they hear. It is often explained by the place of stress or by the acoustically more relevant part of the word for the child. It can be assumed that a third factor might exist, namely the right-to-left fashion of recognition which might result in the storage of the final part of the word. For the same reason, children around the age of 2 normally use many metatheses in their speech which seems to be one more piece of evidence for the regressive direction of word recognition.

Experiments concerning the context-dependency of lexical access

Two experiments were carried out with Hungarian listeners using Hungarian speech materials in order to investigate the context effects in the word recognition process. The first experiment, where the function of age was also taken into consideraton, was devoted to measurements of context effects, while the second one was meant to analyze one further possible type of context effect: the 'time-unit'. In the latter case, our hypothesis was that the time factor as a syllable-size unit might serve as "context" concerning its influence and result.

Experiment I. Method and material

Speech samples were taken out of six male speakers' monologues. All these speakers spoke standard Hungarian as their native language without any speech error. The only difference among the speakers was their speech tempo which ranged from 7.4 sounds/s up to 16.0 sounds/s. Their articulation rate was between 8.3 sounds/s and 17.1 sounds/s. For the speech material of the experiment, 28 words and 28 word combinations have been extracted electrically from these speakers' spontaneous speech by means of the gating-system of the sound spectrograph. The criteria for choosing the speech samples were as follows: (i) there should be both content and function words, (ii) there should be monosyllables as well as two-, three-, four-, and five-syllable words, representing the normal distribution in Hungarian, (iii) there should be words of various grammatical categories and (iv) of various frequencies of occurrence, (v) there should be mainly inflected but also some non-inflected words. There was an effort taken for careful extraction of speech samples, avoiding noises of cutting. The chosen 28 words in "artificial isolation" (in random order) served the speech material for the first set of word lists.

Word combinations consisted of two words, one of them being the target chosen previously for a member of the set of words in isolation. The other word either preceded or followed the target depending on their actual syntactic relationship. 28 word combinations served for the second set of word list for the experiment.

By means of an intensity and fundamental frequency meter, the actual duration of the words and word combinations has been measured. On the basis of the articulation rate of these speech samples three categories were defined, labelled as 'slow/normal', 'accelerated' and 'fast': 5.8–10.4 sounds/s (11 items), 11.4–16 sounds/s (11 items) and 17.3–23 sounds/s (6 items).

First the extracted words in (artificial) isolation, then the sound combinations have been tape recorded (randomly) for the listening/naming tests.

Four groups of listeners took part in the experiments: (i) 4-year-olds, (ii) 5-year-olds, (iii) 6-year-olds (24 subjects each) and (iv) adults (25 subjects).

The adult listeners' task was, after listening to the words, to write down what they had heard, while the children had to repeat them. After 4 weeks they listened to the word combinations.

Results

Table 1 shows the summarized results obtained in the experiments.

Table 1

Subjects	Correct recognition of words (%)	
	in isolation	in combination
4-year-olds	45.5	76.75
5-year-olds	51.75	84.75
6-year-olds	59.8	83.8
adults	66.14	92.85

The results show—not very much surprisingly—that there is a significant difference ($p<0.001$) between the performance of children and adults. On the other hand, there is also a significant difference in the number of instances of correct recognition between the isolated words and the word combinations in all groups. It seems to be the case that contextual effects represent an advantage not only for adults but also for children. Table 2 contains the differences between data obtained for isolated words and words in combinations according to the various articulation rates. It can be clearly seen that the faster the articulation rate of words, the greater the effect of the context.

Table 2

Articulation rate of words (sounds/s)	Correct recognition (%)			
	isolated words		words in combinations	
	children	adults	children	adults
5.8–10.4	80.3	89	94.7	98.75
11.4–16	41.6	62.5	70.4	90.9
17.3–23	20.8	41.3	73.6	85.4

The correct recognition of extracted words depends mainly on their tempo when they appear in isolation. However, for both children and adults, other factors seem to play a definitive role in recognition, like (i) the actual acoustic structure of the word, (ii) the frequency of usage for the subjects, and (iii) the familiarity of the words in the case of children. (i) E.g.: the very low percentage of correct recognition of the word *járt* 'went' in both groups gives evidence for the factor of the acoustic structure. Comparing the data of this word to the average values obtained for its (tempo) group (where their rate is: 5.8–10.4 sounds/s), it is clear that the 25% of its correct identification by children and 44% of correct recognition by adults is the consequence of its acoustic structure. The average score for that group of words is 80.3% in the case of children and 89.0% in the case of adults.

(ii) Some words give evidence for the effect of frequency which is characteristic of the listeners themselves. E.g.: the word *magyar* 'Hungarian' is a very frequent one for the children for some reason, but not so much for adults; however, the word *különböző* 'different' is more frequent for adults than for children. Correct recognition of the word *magyar* which falls into the fast group of words according to its articulation rate, is 58.3% of all cases by children and 72% of all cases by adults. The average score for that group of words is 20.8% in the case of children and 41.3% in the case of adults. Correct recognition of the word *különböző* in the same (tempo) group is 58.3% of all cases by children and 100% of all cases by the adults.

(iii) With the help of the children's kindergarten teachers the familiarity of words to the children has been defined. Nearly 30% of all words were unfamiliar to children in all age groups. In Table 3 results are summmarized.

Table 3

Word groups articulation rate (sounds/s)	Correct recognition of words (%)			
	familiar		unfamiliar	
	isolated	in combination	isolated	in combination
5.8–10.4	81.47	93.5	75	100
11.4–16	41.6	72.6	41.6	66.7
17.3–23	16.65	87.5	29.15	45.8
Average	46.57	84.36	48.58	70.83

Comparing the children's data in the two groups, two consequences, one expected and one only hypothesized, could be drawn: (i) the expected one is that in the case of familiar words the ratio of correct recognition is much better when they appear in a combination; however, there is not as big difference in all groups in the case of unfamiliar words. (ii) The hypothesized process of word recognition for children is that the operations at the lower levels should be more detailed and precise than for adults. This is supported by examining the correct identification data for isolated words which are equal for accelerated words (41.6%) and better for fast words (16.65% and 29.15%) when access to the semantics does not work. It appears that the children focus on acoustic–phonetic structure, i.e. the acoustic and phonetic analysis of the sound sequences they hear. An apparent contradiction might be found in this respect between the correct recognitions of familiar and unfamiliar words with normal articulation rate (5.8–10.4 sounds/s). According to our hypothesis, success rate should have been equal in the two cases since lower-level operations had been predicted to be very precise in this tempo group (particularly for unfamiliar words). Although the difference between the familiar and unfamiliar words is not significant, we suppose that false attempts to access a word (misperceptions or misunderstandings) might have resulted in the incorrect responses. In case of accelerated or fast words there is no possibility for the child to make several trials to find the appropriate item in the mental lexicon.

From another aspect, results have given another piece of support for the more detailed and more precise operations at the lower levels in the case of children: the children repeat exactly what they have heard on the basis of their lower-level decisions—which result in a meaningless sound sequence approaching the perceived one. The adults, on the contrary, try to contact immediately to the lexicon which results in various words because of the uncertain operations at the lower levels and because of the adults' effort to recognize real words.

No one can be surprised to learn that the perception of verbal phrases becomes more accurate with age. It is claimed that this is predictable on the basis of the fact that familiar words are more easily perceived (Maccoby 1971, 24). We think that this fact is only part of an acceptable explanation that should also involve other factors like developing perception base, background knowledge, integration ability of verbal stimuli, etc. The lower levels of the decoding mechanism of children work differently from those of adults: their operations meet the expectations of the whole speech acquisition process. So, children's mechanism is not forced for all cases to contact the acoustic–phonetic data with some representation of the lexicon. They are ready to suppose that there are 'gaps' in their lexicon, and are ready for recognizing newer and newer word-forms prior to understanding the meanings.

Figure 1 summarizes the differences in word recognition according to the place of the target word in the combination (whether the target precedes or follows the other word).

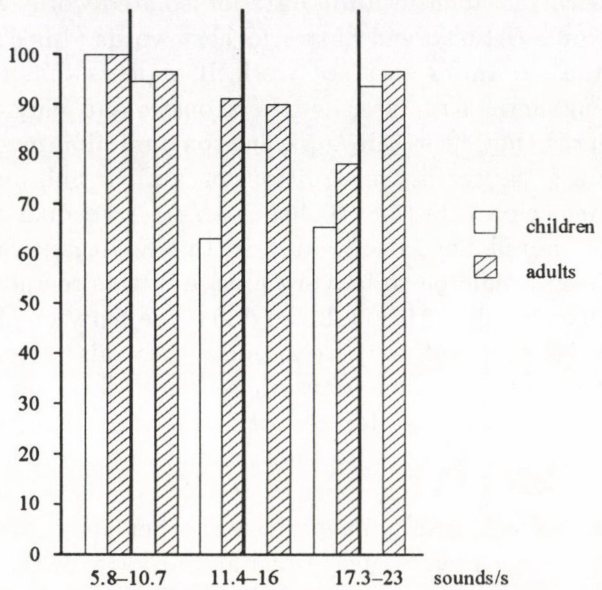


Fig. 1
Effect of the place of the target word in combination on correct recognition (thick lines represent the nontarget word of each combination; the columns show the recognition scores of the target depending on position)

The data of adults show a significant difference only in the case of the very fast words, while the children's data show significant differences for the second and third groups of words (accelerated and fast). It can be claimed that the context effect, whatever it is, appears only when the articulation rate of the target words is faster than 16 sounds/second for adults and it appears when this rate is faster that 11 sounds/second in the case of children. This means that subjects' predictions are much more reliable when the target word follows another word.

We have analyzed the correct recognition of the target words depending on their semantic or syntactic coherence in the word combination. We found no difference in scores either of children or of adults in terms of the target words' being semantically or syntactically related or less related to the other word. However, interesting examples have been found in the material highlighting a specific 'context'-variation for the recognition process. There is the word *tehát* 'consequently' which was correctly recognized in isolation by children in 66.6% of the cases and in 60% of the cases by adults. It was impossible to extract another word to this target to make a word combination, because hesitation preceded and followed it and only one of the Hungarian definite articles appeared close to it which was the [ɔ] vowel. The "word combination" sounded as: *tehát* [ø] (hesitation sound) [ɔ] (definite article), the overall duration was similar to those of other combinations. Surprisingly, subjects performed significantly better when they identified the target word in this funny combination than in isolation; the children's score was: 83.4% and the adults' score was 100%. (All the children imitated also the sound of hesitation.) This is a very clear case which involves neither semantic nor syntactic factors as contextual effects, no linguistic relation at all; however, the contextual effect works.

What is then the cause for the improvement of word recognition in this case? There is only one possible explanation: time might function here as a specific type of context which can be called 'time unit' for word recognition. The second series of experiments was to prove this hypothesis.

Experiment II. Method and material

The aim of this experiment was to find out whether the supposed "time-unit" really works as a specific context in word recognition.

15 Hungarian words have been selected using the same criteria as in the previous experiment. A trained male speaker was asked to utter these words so that after each word an [ø] sound was added (the speaker was explained the situation: to speak the target word fluently and while searching the next item he should use the hesitational sound instead of a silent pause, i.e. the same way as it occurs frequently in spontaneous speech). In this way 15 speech samples were produced that sounded acceptable as "chunks" of natural word combinations where the second part (and always the second one) was missing and substituted by the common Hungarian hesitational vowel. The overall duration of these "word combinations" was similar to those of normal combinations of the previous experiment. The articulation tempo of words ranged between 10.6 sounds/s and 13.7 sounds/s. (The speaker made an effort to speak as naturally as he could.)

Two series of test materials were made: (i) in the first there were the target words in their artificial context with the hesitational phenomenon in random order (e.g. *sok . . . öö* 'many' [ø] /hesitation sound/); (ii) the second material contained only the target words where the originally produced hesitation had been electrically eliminated. Both test materials were masked by white noise (this technique has been known since 1951 when Miller, Heise and Lichten conducted a perceptual experiment with this method). The signal-to-noise ratio was −8 dB in general. A 25-second (noisy) pause was left between the items.

18 adult subjects with a mean age of 35 (females and males) participated in the experiments under two different conditions. In the first test situation half of the subjects listened to the speech material of the target words while the other half of them was provided by the test material with the artificial word combinations. 4 weeks later they heard the other speech material. They were required to write down what they had heard according to a naming task. The experiments were carried out in the same silent room using headphones.

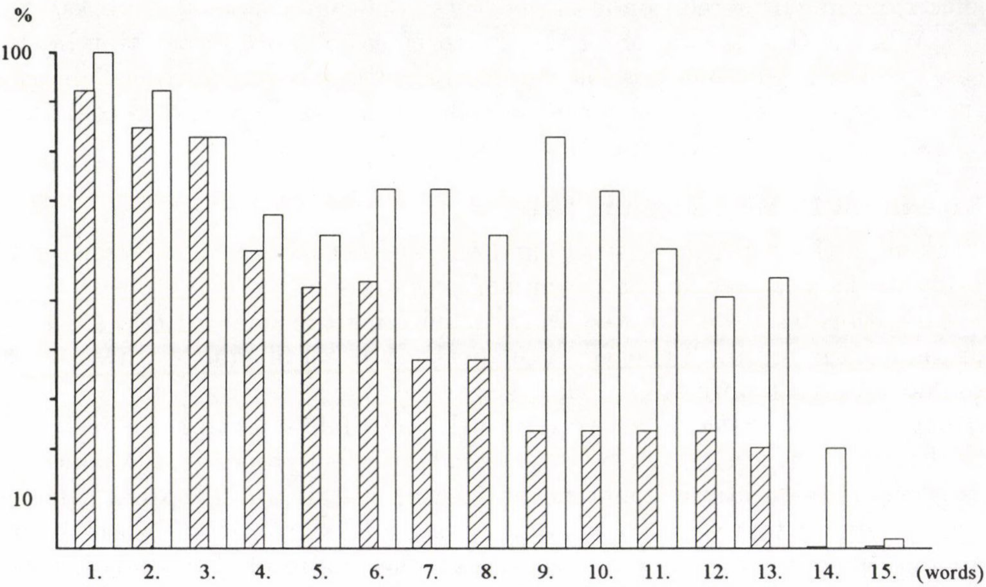


Fig. 2
Correct recognition of words in isolation (▨) and with hesitation sound (□)

Figure 2 shows the correct recognition of target words in the two conditions. It can clearly be seen that the recognition of targets being in "context" was significantly better than in isolation ($p<0.01$). The subjects recognized the total of 15 words correctly in 43.3% of all responses while correct responses were given in 65.5% of all cases when the target words were followed by the hesitational phenomenon. The order of the words (1–15) in

Fig. 2 is as follows: *már* 'alredy', *szaladni* 'to run', *másik* 'another', *róluk* 'about them', *súgott* '(he) whispered', *sok* 'many', *verset* 'poem-Acc.', *utat* 'way-Acc.', *kecske* 'goat', *betűk* 'letters', *amit* 'what-Acc.', *radírt* 'eraser-Acc.', *jöttek* '(they) came', *veszek* '(I) buy', *mezőn* 'in (the) field'.

There are two targets (*mezőn* 'in the field' and *veszek* 'I buy') that the subjects were not able to recognize in isolation (supposedly due to their acoustic structure), however, one of them could be correctly recognized with the hesitation sound in 22.2% of all responses (for the target *veszek* 'I buy'). Surprisingly, phonetic analysis of the responses shows great differences for the other word as well. Since there was almost no response for the word *mezőn* 'in the field' in isolation, various words with similar acoustic structure were written for it occurring in combination (like *mondom*, *majom*, *mellyel*, *mozdul*, *megjön*, *mozdony*). Table 4 shows examples for acoustic and contextual dependency of recognition.

Table 4

Target words		Correct recognition (%)	
IPA	gloss	in isolation	with hesitational phenomenon
ma:r	already	94.4	100
ro:luk	about them	61.1	77.7
ma:ʃik	another	73.3	83.3
ʃu:got:	whispered	55.5	66.6
vesek	I buy	–	22.2
ɔmit	which	24.8	61.1
rɔdi:rt	eraser-Acc.	24.8	52.2

There was no correlation found concerning the frequency of usage of the target words.

Conclusions

The results of this latter experiment have supported our hypothesis that a filled-in pause—in our case the [ø] hesitation—functions as context for word recognition. This means that the hesitational phenomenon fulfils the same role in the recognition process as another word does being in either semantic or syntactic relation (or both) with the target. Within the framework of the analysis-by-synthesis model (cf. Klatt 1980) this specific function of hesitation can easily be explained. If we think of the effect of context we have to reckon with the factor of time which—as the data show—can replace the syntactic/semantic context effects. We can assume that if the listener is provided with more time than it is usual in the recognition process of isolated words, the perceptual mechanism is set up for operations containing the elaborations of context, too. This should be the case even if the context is lost (for some reason) or replaced as in our experimental case. The word recognition processes activated are assumed to expect also contextual effects. During hesitation, lexical access will be similar to that in the case of word combinations, i.e. better and/or more detailed operations can take place at each level of the mechanism. Feedback phenomena can influence the supposed (recognized) word since the recognition process of the hesitational sound requires less activity from the mechanism, so it might still focus on the previous word.

Many questions arise concerning these results, first of all the language-specificity of contextual effects, that are worth analyzing further.

References

Altmann, G.T.M. 1991. Cognitive Models of Speech Processing. MIT Press, Cambridge, MA.

Broadbent, D.E. 1967. Word-frequency effect and response bias. In: Psychological Review 74: 1–15.

Cole, R.A.–Jakimik, J. 1980. A model of speech perception. In: Cole, R.A. (ed.): Perception and Production of Fluent Speech, 133–60. Erlbaum, Hillsdale, NJ.

Cowan, N.–Leavitt, L.A.–Massaro, P.W.– Kent, R.D. 1982. A fluent backward talker. In: Journal of Speech and Hearing Research 25: 48–53.

Cutler, A. 1986. *Forbear* is a homophone: lexical prosody does not constrain lexical access. In: Language and Speech 29: 201–21.

Cutler, A.–Butterfield, S. 1990. Durational cues to word boundaries in clear speech. In: Speech Communication 9: 485–95.

Cutting, J.E.–Pisoni, D.B.–Cole, R.A.–Jakimik, J. 1980. A model of speech perception. In: Cole, R.A. (ed.): Perception and Production of Fluent Speech, 133–63. Erlbaum, Hillsdale, N.J.

Elliot, L.L.–Clifton, L.A.–Servi, D.G. 1983. Word frequency effects for a closed-set word identification task. In: Audiology 22: 229–40.

Feldman, L.B. 1991. The contribution of morphology to word recognition. In: Studdert-Kennedy, M. (ed.): Report on Speech Research of Haskins Lab., 231–43.

Forster, K.I. 1979. Levels of processing and the structure of the language processor. In: Cooper, W.E.– Walker, E. (eds): Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett. Erlbaum, Hillsdale, N.J.

Forster, K.I. 1989. Basic issues in lexical processing. In: Marslen-Wilson, W.D. (ed.): Lexical Representation and Process, 75–108. MIT Press–A Badfork Book, Cambridge, MA.–London.

Frauenfelder, U.H.– Komisarjevski-Tyler, L. (eds): 1987. Spoken Word Recognition. MIT Press, Cambridge, MA.–London.

Frauenfelder, U.H.-Lahiri, A. 1989. Understanding words and word recognition: Does phonology help? In: Marslen-Wilson, W.D. (ed.): Lexical Representation and Process, 319-42. MIT Press-A Badfork Book, Cambridge, MA.-London.

Frauenfelder, U.H.-Segui, J.-Dijkstra, T. 1990. Lexical effects in phoneme processing: Facilitatory or inhibitory? In: Journal of Experimental Psychology: Human Perception and Performance 16: 77-91.

Fry, D.B.-Denes, P. 1952. Mechanical speech recognition. In: Jackson, W. (ed.): Communication Theory, 426-32. London.

Gósy, M. 1987. High frequency speech perception: Phonetic aspects and application. Proceedings of XIth ICPhS. Vol. 1, 88-92. Tallinn.

Gósy, M. 1992. Speech Perception. Hector, Frankfurt.

Johnson-Laird, P.N. 1987. The mental representation of the meanings of words. In: Cognition 25: 189-211.

Kintsch, W.-Mross, E.F. 1985. Context effects in word identification. In: Journal of Memory and Language 24/3: 336-50.

Klatt, D. 1980. Speech perception: A model of acoustic-phonetic analysis and lexical access. In: Cole, R.A. (ed.): Perception and Production of Fluent Speech, 243-89. Erlbaum, Hillsdale, N.J.

Klatt, D. 1989. Review of selected models of speech perception. In: Marslen-Wilson, W.D. (ed.): Lexical Representation and Process, 169-227. MIT Press-A Badfork Book, Cambridge, MA.-London.

Ladefoged, P. 1967. Three Areas of Experimental Phonetics. London.

Maccoby, E.E. 1971. Age changes in the selective perception of verbal materials. In: Morton, D.L.-Jenkins, J.J. (eds): The Perception of Language, 23-4. Charles E. Merril Publishing House, Columbus, Ohio.

Marslen-Wilson, W.D. 1989. Access and integration: Projecting sound onto meaning. In: Marslen-Wilson, W.D. (ed.): Lexical Representation and Process, 3-25. MIT Press-A Badfork Book, Cambridge, MA.-London.

McClelland, J.L.-Elman, J.L. 1986. The TRACE model of speech perception. In: Cognitive Psychology 18: 1-86.

Miller, G.A.-Heise, G.A.-Lichten, W. 1951. The intelligibility of speech as a function of the context of the test materials. In: Journal of Experimental Psychology 41: 329-35.

Morton, J. 1969. Interaction of information in word recognition. In: Psychological Review 76: 165-78.

Nooteboom, S.G.-Truin, P.G.M. 1980. Word recognition from fragments of spoken words by native and non-native listeners. In: IPO, Annual Progress Report 15: 42-7.

Pisoni, D.B.-Luce, P.A. 1987. Acoustic-phonetic representation in word recognition. In: Frauenfelder, U.H.-Komisarjevski-Tyler, L. (eds): Spoken Word Recognition. MIT Press, Cambridge, MA.-London.

Samuel, A.G. 1986. The role of the lexicon in speech perception. In: E.C. Schurab-Nusbaum, H.C. (eds): Pattern Recognition by Humans and Machines: Speech Perception. Vol. 1. Academic Press, New York.

Schreuder, R.-G.F. D'Arcais. 1989. Psycholinguisticissues in the lexical representation of meaning. In: Marslen-Wilson, W.D. (ed.): Lexical Representation and Process, 409-37. MIT Press-A Badfork Book, Cambridge, MA.-London.

Seidenberg, M.S. 1985. The time course of information activation and utilization in visual word recognition. In: Besner, D.T.–Waller, T.G.–MacKinnon, G.E. (eds): Reading Research: Advances in Theory and Practice. Vol. 5. Academic Press, New York.

Seidenberg, M.S.–Tanenhaus, M.K.–Leiman, J.M.–Bienkowski, M. 1982. Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. In: Cognitive Psychology 14: 489–537.

Slobin, D.I. 1970. Universals of grammatical development of children. In: d'Arcais, G.F.–Levelt, J. (eds): Advances in Psycholinguistics, 174–186. Foris Publications, Amsterdam.

Slowiaczek, L.M. 1990. Effects of lexical stress in auditory word recognition. In: Language and Speech 33: 47–68.

Solomon, R.–Postman, L. 1952. Frequency of usage as a determinant of recognition thresholds for words. In: Journal of Experimental Psychology 43: 195–201.

Stevens, K.N. 1978. The speech signal. In: Kavanagh, J.F.–Strange, W. (eds): Speech and Language in the Laboratory, School, and Clinic, 3–28. Cambridge, MA.–London.

Tanenhaus, M.K.–Lucas, M.M. 1987. Context effects in lexical processing. In: Frauenfelder, U.H.–Komisarjevski-Tyler, L. (eds): Spoken Word Recognition, 213–35. MIT Press, Cambridge, MA.–London.

Zwitserlood, C.M.E. 1989. Words and Sentences: The Effects of Sentential-Semantic Context on Spoken-Word Processing. Max Planck Institute for Psycholinguistics, Nijmegen.

Address of the author: Mária Gósy
Research Institute for Linguistics
Hungarian Academy of Sciences
Színház u. 5–7.
1250 Budapest, P.O. Box 19.
Hungary

Acta Linguistica Hungarica, Vol. 42 (1–2), pp. 51–62 (1994)

SOUND DURATION MEASUREMENTS IN DECLARATIVE SENTENCES*

GÁBOR OLASZY

Introduction

The study of the temporal structure of speech and of speech sounds is one of the most traditional fields in speech research. In spite of this, many questions cannot be answered even today, especially those concerning the rhythmic structure of fluent speech. In Hungary the first timing measurements were done at the beginning of the 20th century by Gombocz (1909). The latest monograph was written by Kassai (1979), who summarised the history of time measurements that have been done in Hungary and also published the results of her own investigations about the duration of Hungarian speech sounds. Further time measurements were carried out, using the analysis by synthesis method, from the early 80s (Olaszy 1985; 1989).

New demands have arisen with respect to time structure measurements in speech; in particular, ones concerning the question of temporal variation in continuous, fluent speech. This type of demand became especially pressing when the first computer generated text-to-speech systems were constructed. Automatic speech synthesis from written texts produced monotonous, rhythmless but understandable speech signals. So it became clear that to devise the right time structure and its continuous variation during the production of longer texts (sentences) is one of the most important elements in speech synthesis. René Collier characterised this problem in 1992 as follws: "At present, several text-to-speech systems produce fairly intelligible output at sentence level, but they lack naturalness, individuality and expressiveness. It is unlikely that an inexperienced listener would be able to listen to long passages of running artificial speech and comprehend the message correctly with little mental effort" (Collier *et al.* 1992, 32). He emphasises in this paper that a lot of research has to be done in the future for reaching a more natural voice quality in artificially generated speech. The fields of time structure variations

* This research was supported by OTKA [Project 3211].

and melody patterns are to be concentrated on. Concerning Hungarian, some preliminary research has been done and rules based on it were built into the MULTIVOX multilingual text-to-speech system. Such rules were a shortening rule used for articles (*a*, *az* 'the'), its opposite, a lengthening rule used for one-syllable question words (*Mi?* 'What?', *Ki?* 'Who?', *Én?* 'Me?', *Most?* 'Now?') and a joining rule to organise certain words into one fluent sequence (*Nemtudom* 'I don't know', *Senkisem* 'Nobody does', *Hogyvagy?* 'How are you?', *Mikorindulunk* 'When do we start?') i.e. to implement continuous articulation on the segmental level (Olaszy 1989). These simple rules resulted in a better speech quality but it was still far from human speakers' rhythm and intonation. This negative result can be explained by the fact that in this system the elements of each sound combination are always the same (in frequency and time structure) independent of their position in the word and independent of the place of the word in the sentence. Hence, the time and frequency structure of the synthesized sentence would remain too regular.

The aim of the research

1. We search for rules in the time structure of speech which are governed by the language itself and which determine the duration of sounds, sound combinations and longer sequences.

2. We would like to characterise temporal structure in terms of various suprasegmental elements with exact data and rules.

3. The results are to be used in automatic speech generation and in speech recognition.

The linguistic corpus

The material for this research contained declarative sentences (news items) picked out of newspapers. 20 sentences (of various length) were read out by a male speaker and were recorded on tape.

The characteristic data of these sentences were:

number of words per sentence	4–8
number of syllables per sentence	14–27
number of sounds per sentence	32–53

The total number of analysed words was 105, that of syllables was 344 and that of sounds was 760. The announcer's average fundamental frequency was 117 Hz. Some examples from the sentences: *Pozsgai Imre lemond a rádió* *és televízió felügyeletéről.* 'Imre Pozsgai resigns as the coordinator of the radio and television'. *Pozsgai Imre az alábbi nyilatkozatot juttatta el a távirati irodához.* 'Imre Pozsgai declared the following through the Hungarian News Agency.' *A Minisztertanács rám bízta a rádió és televízió felügyeleti jogát.* 'The Council of Ministers granted me the right of supervision of the radio and television.'

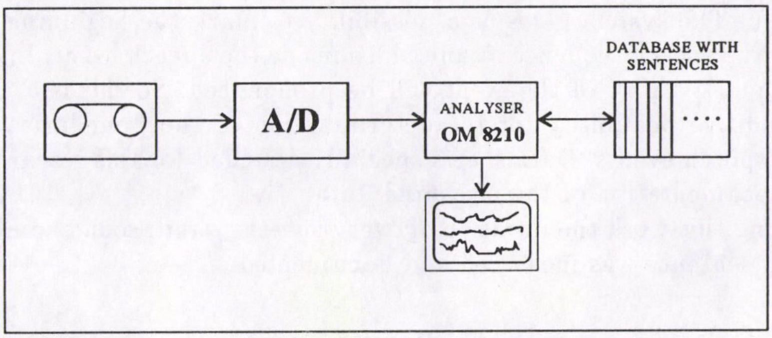


Fig. 1

Method

1. Creating a data base

We have started to compile a sentence data base where the acoustic data of the examined sentences are stored in the form of 12 ms long speech frames (every frame includes 5 formant values, 5 formant bandwidths, amplitude, pitch, and source information). The Philips OM 8210 development system (Zelle–Roberts 1988) was used for the preprocessing of sentences and for saving them in a file (Fig. 1). This system gives us the possibility to get visual data from the inherent acoustic structure of the given sentence in phonetic terms and in the meantime to perform auditive checking of the sentence or parts of it (even of one speech frame). The system performs speech data compression and transformation; so the result is a phonetic–acoustic representation of the given sentence. When resynthesising the sentence into analog form the sounding will be very similar to the original one. Sentences in such representation constitute the elements of the data base. Every sentence can be accessed at any time and listened to or used for various purposes (e.g. duration measurements, pitch determination etc.).

2. Time structure measurements

The duration measurements were done manually as follows: the sentence was loaded from the data base, the acoustic parameters appeared on the screen (in a form similar to a sound spectrogram). The audio output function of the editor was switched on so that the acoustic content could be listened to by a step forward–backward rolling procedure frame by frame. So the sound boundaries could be determined on the basis of the screen data and on that of auditive perception. The system gives us a possibility to mark the beginning and end points of the speech sequence at any place inside the speech event. In this case only the marked part of the event will be pronounced. So this is a combined visual–auditive possibility for the determination of the boundaries between different speech events. Data sheets specially designed for this research served for the documentation of the measured data.

During this work the duration of every speech sound, sound combination, word and sentence was measured and documented.

Results

After the processing of the 20 sentences, the following results can be summarised.

Time structure on sentence level

The speech wave is continuous inside sentences: no measureable breaks suggesting word boundaries can be seen between words, i.e. sentence level articulation is fully continuous. The word level information to demarcate words from one another in the continuous speech event is found first of all in the pitch structure. Even in special cases, when identical sounds meet at a word boundary, this phenomenon is observable. For example, in the sequence *Pozsgai Imre* ... the final [i] of the word *Pozsgai* and the initial [i] of the word *Imre* show no change in their energy curve (the sounds are fully amalgamated and are represented as a long [i:]). But there is a pitch fall in the word final [i] and a pitch rise in the beginning position [i] as a representation of the word boundary between the two words. This is the realisation of first-syllable accentuation in words. A similar case can be found in the sequence ... *rám bízta a rádió* ... in the [ɔ] vowels.

In both cases a minimum of 10 Hz rise can be measured in the second half of the resulting long vowel as a marker of word boundary. A further observation can be made with respect to the second case where a final [ɔ] sound and a definite article [ɔ] are joined to each other. Measurements concerning the pitch structure in articles show that the pitch in the article is generally lower or the same (but not higher) as in neighbouring words. So in this case we would also expect lower pitch in the article than in the preceding and following words. The special situation (i.e. joining together the same vowel at the word boundary) brings about a special structure, too. The measured data show that the pitch is 10 Hz lower in the final [ɔ] sound in the word *rám bízta* (90 Hz) than in the following article (100 Hz) and only in the next word *rádió* does the pitch rise 15 Hz higher after the article (115 Hz). So in this special case our word boundary marker mechanism sets the pitch lower before the article than in the article itself to ensure the pitch rise phenomenon as a word boundary marker in the article.

Time structure on word level

Two categories of words were measured: the final word of the sentence on the one hand and all the other words independent of their position in the sentence on the other.

Word duration shows a linear rise as a function of the number of syllables (Table 1). Final words in the sentence are always longer than the others.

Table 1

Number of sounds in the word	1	2	3	4	5	6	7	8	9	10	11	12	13
	The number of relative duration units to express the length of the word												
Duration of the last word	–	–	32	41	46.5	56	–	74.6	77.6	92	–	104	–
Duration of all other words	7.6	14.8	22.3	27.2	35.3	43.8	51.6	55.6	65	67.4	69.6	84.3	–

The numbers in the table express the duration of words in relative duration units (1 unit is 8.8 ms long). This means that, for example, the duration of a word containing three sounds is defined as 22.3 units, i.e. the physical duration is 22.3×8.8 ms = 196 ms. The same three-sound word in a final position has a duration of 32 units, i.e. 281 ms (Fig. 2).

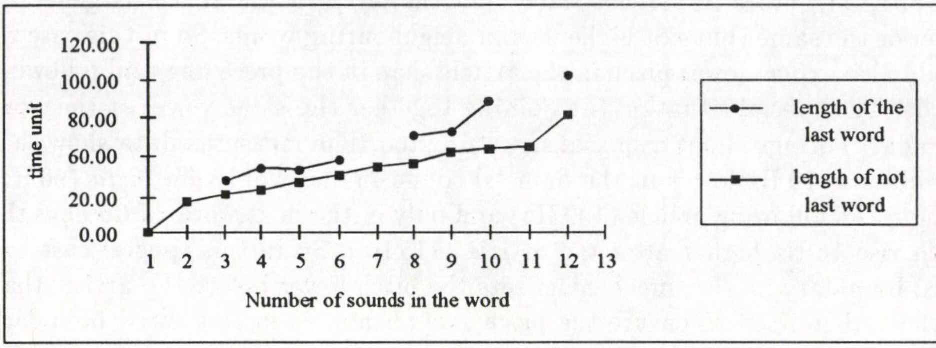


Fig. 2

The diagram and the data show that the position of a word inside the sentence has no influence on the relative duration of the given word. Only sentence final words have longer duration. The average duration of words in the examined sentences is as follows:

words of 1 sound	(average of 16 words)	67 ms
words of 2 sounds	(average of 9 words)	130 ms
words of 3 sounds	(average of 10 words)	196 ms
as a last word		281 ms
words of 4 sounds	(average of 5 words)	239 ms
as a last word	(average of 2 words)	360 ms
words of 5 sounds	(average of 6 words)	310 ms
as a last word	(average of 2 words)	409 ms
words of 6 sounds	(average of 14 words)	385 ms
as a last word		492 ms
words of 7 sounds	(average of 5 words)	454 ms
words of 8 sounds	(average of 6 words)	489 ms
as a last word	(average of 3 words)	656 ms
words of 9 sounds	(average of 9 words)	572 ms
as a last word	(average of 3 words)	682 ms
words of 10 sounds	(average of 7 words)	593 ms
as a last word	(average of 2 words)	809 ms
words of 11 sounds	(average of 5 words)	612 ms
words of 12 sounds	(average of 3 words)	741 ms
as a last word	(average of 2 words)	915 ms
words of 14 sounds	(average of 3 words)	838 ms
as a last word		1020 ms
words of 15 sounds	(average of 2 words)	840 ms
as a last word		932 ms

On the basis of the above data we can predict the duration of a word for speech synthesis applications on the sentence level.

The tempo of articulation was measured on word level, too. For this examination six sentences, each containing 7 words, were selected. The number of syllables in these sentences was between 21–23, the number of speech sounds was between 46 and 50. So the length of the sentences was very close to equal. The examination was based on the duration value of each speech sound in the given word. The results are summarised in Table 2, the tendencies concerning word tempo are shown in Fig. 3.

Table 2

The examined sentences	1st word	2nd word	3rd word	4th word	5th word	6th word	7th word
	relative duration units/sound						
1st sentence	8.8	8.7	6.3	6.2	8.5	6	7
2nd sentence	7.1	7	7.5	9	6	8.7	8.8
3rd sentence	7	7	6.6	7.7	7.5	7.6	8.7
4th sentence	6.7	6.9	6.6	6.8	7.6	8.2	10.3
5th sentence	7	6.2	6.3	7	12	7.2	8.8
6th sentence	8.8	6	8.3	7.5	7	9.6	9.1

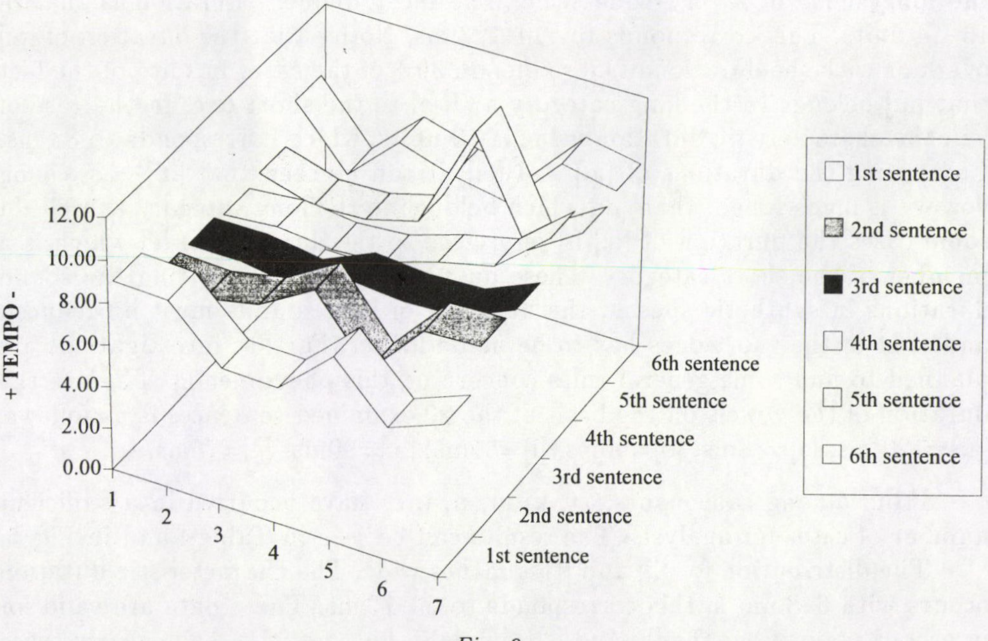


Fig. 3
The articulation tempo of the words in the six sentences

The data in Table 2 are to be read as follows: faster tempo=smaller number, slower tempo=larger number. The data express that the word tempo is faster inside the sentence (in words 3,4,5) than at the beginning or at the end. Of course this can be treated only as a general rule. There were exceptions where the tempo was slower than predicted. Examining these words, we found that most of them were emphasised. For example in the 4th word of the 2nd sentence the measured tempo was very slow (value=9). This word was pronounced with a great emphasis.

Duration of sounds

A total of 304 sounds were measured in the 20 sentences. Of course, some sounds did not occur in large enough numbers for an average to be calculated. The durational data concerning vowels are shown in Table 3 and in Fig. 4.

It can be seen that the most homogeneous duration is exhibited by the sound [e]. The maximum of the distribution is represented by the duration units 7,8,9,10. This corresponds to a duration between 61 and 88 ms. 77% of [e] sounds belongs to this category. The [a:], [ɔ], [e:] and [i] sounds show a less homogeneous distribution in duration. [a:] and [e:] belong to the category of long vowels in Hungarian. In the case of [a:], the shortest duration is 7 units, the longest is 20: a very wide scale. The most frequent duration is that of 13–14 units. This corresponds to 114–123 ms. Notice that the duration of [a:] overlaps with the duration of [ɔ] sounds in 30% of the cases in spite of the fact that [a:] belongs to the long category and [ɔ] to the short one. In the case of [e:], the characteristic duration value is 10 units, which corresponds to 88 ms. Comparing the durations of [a:] and [e:], it can be seen that [a:]—as a long vowel—is much longer than [e:] which belongs to the long category as well. In some cases the duration of [e:] is very close to the duration of [e] which is a member of the short category. These data suggest that in setting the sound durations in synthetic speech, the duration of long sounds must be reduced and that of the short ones has to be made longer. Further investigations are planned to find some general rules concerning this phenomenon. The average duration of the vowels on the basis of the 20 examined sentences is as follows: [a:]=120 ms, [ɔ]=80ms, [o]=90ms, [i]=55ms, [e:]=90ms, [e]=75ms.

From among consonants, /t, k, m, n, l, r/ have occurred in a sufficient number of cases for analysis. The results can be seen in Table 4 and in Fig. 5.

The distribution for [k] and [t] is rather wide. The characteristic duration occurs with 6–8 units, this corresponds to 53–70 ms. These data are valid for intervocalic position; the [k] and [t] sounds in final position have nearly twice as long durations.

Table 3

	The number of relative duration units in the sound													
	3	4	5	6	7	8	9	10	11	12	13	14	15	16
sound	the number of occurrence in the given duration													
[ɔ]	1	2	4	5	10	8	14	6	4	1	1			
[a:]					1	1	1	3	4	3	7	8	3	
[o]			2	2	4	3	4	7	2	1				
[i]	1	1	6	12	10	5	4	6	1			2		
[e]		1		8	17	20	19	17	6	4		1		
[e:]				1			1	8		4		1	1	1

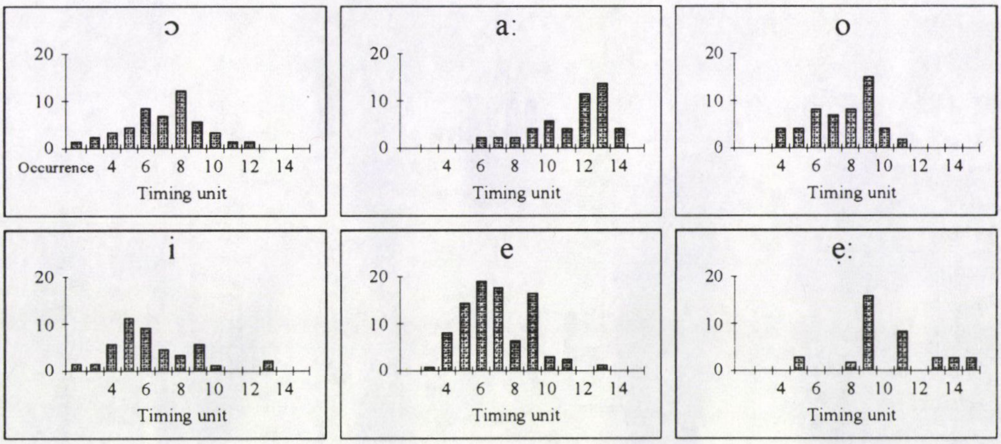


Fig. 4
The distribution of durations of vowels, expressed in terms relative duration units

The duration data for nasals—[m] and [n]—suggest the following distinction between the two sounds. With respect to [m], the maximum of the distribution is at 4–6 units, i.e. 32–52 ms, while for [n] this value is at 3 units, i.e. 26 ms. So there is a characteristic difference between the typical durations of these two sounds. That fact can be important in speech synthesis where the proper generation of these two sounds is always problematic. In most synthesis systems the acoustic structures of these two sounds are very close to each other because of the limitations of the synthesis method or the synthesizer. Therefore, this duration distinction can prove to be important for the characterisation of these two sounds in fluent synthetic speech.

Table 4

	The number of relative duration units in the sound													
	2	3	4	5	6	7	8	9	10	11	12	13	14	15
sound	the number of occurrence in the given duration													
[k]			3	5	7	4	7	5	2		3			
[t]			3	10	17	9	14	5	5	4	3			
[m]		3	5	7	6	2								
[n]	6	11	1	7	2	2								
[l]	10	23	8	5	2	1	3							
[r]	5	17	5	2	3	2								

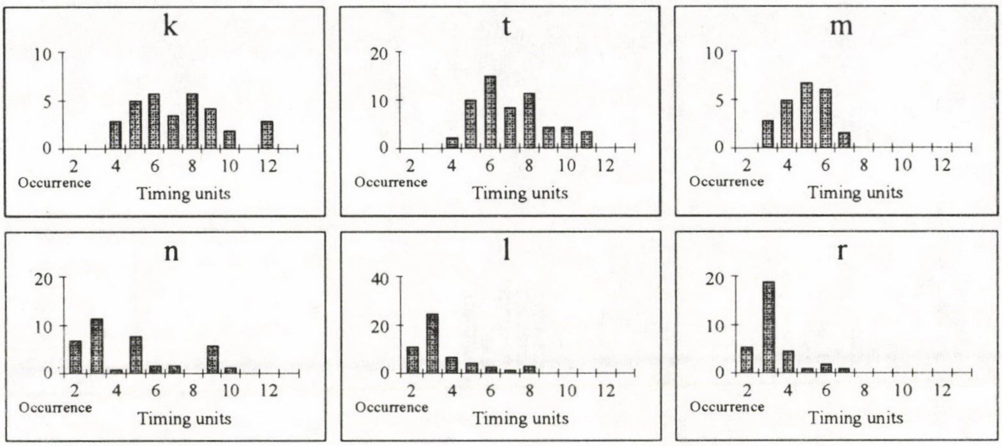


Fig. 5

The distribution of duration of consonants, expressed in terms of relative duration units

For the [l] and [r] sounds the characteristic duration is at 3 units, i.e. 26 ms. This value could be predicted for [r] but is very short for [l]. Earlier measurements showed nearly twice as long duration for [l].

Comparing the measured data with earlier measurements

The above results were compared with those of three Hungarian researchers (Magdics 1968; Szende 1976; Kassai 1978). An important fact—which is also discussed by Kassai in this volume—is that the speech tempo of Hungarian has increased. The most radical shortenings concern the following sounds:

sound	shortening in %
[a:]	15
[e]	25
[k], [t]	40
[m]	45
[n]	55
[l]	35

Discussion

Concerning the temporal structure of simple declarative sentences, the following general conclusions can be drawn from these investigations.

1. The duration of the vowels can be predicted for synthetic speech generation (for a given speed).

2. The duration of words can be predicted, so the proper adjustment of word durations can help to get closer to the original rhythm of fluent speech.

3. At the beginning and at the end of the sentence a speed control factor can be given for the adjustment of a slower speech tempo.

References

Collier, R.–Leeuwen, H.C.–Willems, L. 1992. Speech synthesis today and tomorrow. Philips Journal of Research 47: 15–34.

Gombocz, Z. 1909. A magyar hangok időtartamáról [About the Duration of Hungarian Speech Sounds]. In: Nyelvtudomány II (Linguistics Vol. 2), 93–100. Budapest.

Kassai, I. 1979. Időtartam és kvantitás a magyar nyelvben [Duration and Quantity in Hungarian]. In: Nyelvtudományi Értekezések 102. Akadémiai Kiadó, Budapest.

Magdics, K. 1969. A magyar beszédhangok időtartama [The Duration of Hungarian Speech Sounds]. In: Nyelvtudományi Értekezések 67. Akadémiai Kiadó, Budapest.

Olaszy, G. 1985. A magyar beszéd leggyakoribb hangsorépítő elemeinek szerkezete és szintézise [The Structure and Synthesis of the most Frequent Sound Sequences in Hungarian]. In: Nyelvtudományi Értekezések 121. Akadémiai Kiadó, Budapest.

Olaszy, G. 1989. Elektronikus beszédelőállítás. A magyar beszéd akusztikája és formánsszintézise. [Electronic Speech Generation; The Acoustics and Formant Synthesis of Hungarian Speech]. Műszaki Könyvkiadó, Budapest.

Olaszy, G.–Németh, G.–Gordos, G. 1992. The MULTIVOX multilingual text-to-speech converter. In: Bailly, G.–Benoit, C.–Sawallis, T.R. (eds): Talking Machines: Theories, Models and Designs, 385–411. Elsevier, Amsterdam.

Szende, T. 1976. A beszédfolyamat alaptényezői [Fundamentals of the Speech Process]. Akadémiai Kiadó, Budapest.

Zelle, H.–Roberts, M. 1988. OM 8210 Users Manual. Philips, Eindhoven.

Address of the author: Gábor Olaszy
Research Institute for Linguistics
Hungarian Academy of Sciences
Színház u. 5–7.
1250 Budapest, P.O. Box 19.
Hungary

Acta Linguistica Hungarica, Vol. 42 (1–2), pp. 63–74 (1994)

DO HISTORICAL CHANGES REPEAT THEMSELVES? (ON HISTORICAL 'TWO-OPEN-SYLLABLE SHORTENING' AND PRESENT-DAY 'FAST-SPEECH SYLLABLE ELISION' IN HUNGARIAN)

TAMÁS SZENDE

In this paper two phonological processes will be compared.

(i) The historical tendency of 'Two-Open-Syllable Shortening' (TOSS), first identified and described in detail by Antal Horger as early as in 1911, was a very productive process in Old Hungarian active mainly between the 10th century and the middle of the 14th century bringing about new paradigms both in the declension and the conjugation systems by syncopation, cf.

[serelemV]	'love'	⟶	[serelmet]	'love-ACC'
[urusagV]	'land'	⟶	[orsa:g]	'land'
[mevetegV]	'laugh-IMP'	⟶⟶	[neveʃ:]	'id.',

etc., see below.

(ii) The opposite number of TOSS in this comparison is a present-day postlexical (lenition) process which I will—somewhat loosely—label by a cover term as 'Fast-Speech Syllable Elision' (FSSE) (for details, see point **2**, below).

Since the two process types result in quite similarly structured, in some cases indeed in completely homomorphous, surface forms and TOSS and FSSE are distanced from one another by some eight centuries, the question in the title may appear to be forced. Furthermore, also theoretical difficulties arise concerning this kind of approach. Namely, for example, even if we consider one and the same language, structural differences pertaining to two historical periods split this language into two morphologically distinct, though closely related, variants. A typological/universalistic way of looking at things, however, legitimates such an approximation *per se*. And as for the former question, in what follows I will also try to suggest that "similar", more precisely homophonous, forms show a number of categorical differences. In sum, what will be done in this study is to explore facts and counterfacts in the borderline

1216–8076/94/$ 4.00

area of the 'synchronous' and the 'diachronic' planes of language, to put it in the classical terminology of Saussure (1916/1968, esp. 177–179). In so doing, surface phenomena will serve as the point of departure.

1. The main claims of TOSS (also called Horger's Law in the Hungarian literature) may be summed up on the basis of Horger (1911), Melich (1910), Pais (1952), and Bárczi (1954/1958) as follows.

(i) In a word form consisting of three or more than three syllables, including at least two open syllables, the [–long] vowel of the second (sometimes the third) light, i.e. open and short, syllable may be systematically deleted. Concerning the number of the syllables we should add the following constraint to this rule: Old Hungarian word final reduced vowels prove to be non-syllabic. In view of the data, however, this claim appears at first sight to be too categorical. Namely, holding this view, considerable difficulties arise in the derivation of *Fadd* <geographical name> from *fotudi* in TihAl. [The Deed of Endowment of the Abbey of Tihany issued in Latin by King Andreas I in 1055] and *πaτaδί* in VeszprAl. [The Deed of Gift of the Nuns in the Veszprém Valley issued in Greek by King Stephanus I, also called St. Stephan, before 1002, extant in the form of a *renovatio* from 1109], with the latter virtually including word final reduced [ɪ]. To resolve this problem two possible ways of explanation are available, both having equal chance of motivation. (a) The input of the process should have been in fact one of the casus obliquus forms of the word (such as, for example, that with a *-t* (ACC) formative) in which case the two-open-syllable criterion is met. (b) The second assumption is that word final *-i* still stands for a (non-reduced) [i] in the *-ti/-di* formative in this period of time. Accordingly, the syllabic pattern of the word form here is postulated to have been monotonous with three light syllables. Both (a) and (b) can be maintained. The explanation under paragraph (i), above, is obvious without further ado.

The second path of explanation can be argued for as follows. The text of the TihAl. contains a total of eight (or perhaps nine) Hungarian entries with the variants of the formative *-ti* or *-di* indicating a [–lab +high] vowel word finally, irrespective of whether the preceding segments—both vowels and consonants—involve one or more [+lab] components (see *Opoudi* <geographical name> with a [+lab] component in all segments vs. *ſegis·ti* <geographical name> 'the Lake S.' being [–lab] in all its segments). In all other word final positions a [+lab +high] vowel appears, again, irrespective of the vowel's class membership in the dimension of the [LAB] constituent. These data uniformly indicate the presence of a reduced vowel at this position of the sequence. This is indeed consistent with the predictions of Losonczi's Law according to which the historical process of reduction in stem-final vowels at least in some parts of the Hungarian speaking community must have gone through a parallel labialisation process as well (cf. Losonczi 1923–27, 402 ff). (It is unknown whether the exceptions in the text: *eri iturea*, *fidemſi* and perhaps also *aſauuagi* are reflexes of dialectal differences. The rest with -*C#* and -*CV#* such as *kut*, *ziget*, *hotolm*, and *urusag* are irrelevant in this respect.) This order of things and also the fact that personal and geographical names occupy a distinguished status in (mor)phonology, render -*i* in *fotudi* probable to be a full, i.e. non-reduced, vowel.

(ii) The elision of vowels can be assumed to have passed the intermediate phase of reduction, cf., for example, *ur*[^{u}]*zag* ⟶ *ország* 'land'.

(iii) Vowel reduction was occasionally accompanied by compensatory lengthening, cf. *palica* ⟶ *pálca* 'stick', *capitulum* ⟶ *káptalan* 'chapter [of a cathedral]', or $ari\breve{q}_{[\mathrm{back}]}$. ⟶ *árok* 'ditch'.

(iv) Every vowel can be subjected to TOSS, still "the most frequent cases are those of high and mid vowels" (see Bárczi 1954/1958, 59).

(v) We should add to the foregoing that in the position of the first light syllable within a pair of word boundaries a closed syllable containing a [+cont] consonant may also occur in Old Hungarian, see e.g. *Stephanus*, i.e. /$V_{[\mathrm{front}]}$ʃtefa(:)n/ ⟶ *István* <personal name>, $t\tilde{o}ltV_{[+\mathrm{front}]}s\acute{e}r$, ⟶ *tõlcsér* 'funnel' (if we accept the etymological interpretation of *tõlcsér* as *tõlt-* <verb stem> + /-J(-)/$_{\mathrm{IMP}}$ + -*ér* <formative>).

(vi) As for historical aspects, the domain of the productivity of TOSS extends over centuries: (vi/a) in Proto- and Old Hungarian a number of substantives and whole paradigms of conjugation were systematically involved, cf. HB (= *Sermo Super Sepulchrum*) (around the end of the 12th century, ?1180): *turchucat* 'their throats-ACC', *hotolm* 'power' and IMPSg3 $warV_{[\mathrm{back}]} + \dot{g}V_{[\mathrm{back}]} + n$ ⟶ *várjon* '(s)he should wait' or $mevetV_{[\mathrm{front}]} + \dot{g}V_{[\mathrm{front}]} + n$ ⟶ *nevessen* '(s)he should laugh'; (vi/b) during the period between the 12th and 14th centuries it remained active without any paradigmatic and (mor)phonological restrictions; and (vi/c) after that it has sporadically affected some items only, such as *gyer(e)kõc* 'kid' and *az(u)tán* 'then'.

(vii) In his survey of the issue Bárczi (1954/1958, 60) suggests concerning the phonetic mechanism of TOSS that "[t]he elision [of the vowel] is obviously due to stress [placement]. Namely, as a rule, [the placement of] dynamic stress is followed by a considerable decrease of intensity which the short vowel will be a victim of". In other words: the *punctum saliens* of the change resulting from TOSS is the effect the relative difference existing between the intensity values of stressed and unstressed syllables produces on the architectonics of the word form.

2. The opposite number of TOSS in this comparison, FSSE—a systematically occurring distortion process in present-day Hungarian casual speech—will here be briefly discussed with a syntactic–semantic limitation of the occurrences. Since in this style of speech communication so-called form words may simultaneously go through a number of lenition process types extremely different in nature (such as apocope, reduction, and sequence size truncation, cf. Szende 1992) and, as a consequence, the output may—and as a matter of fact also will—be distorted in a complex, multitudinous way, the comparison will be restricted in this area of occurrences to pure types, i.e. to those of virtual predicates. The conditions motivating syllable elision in content words—ignoring indirect, non-phonetic circumstances such as effects resulting from far-reaching contextual attributes or non-linguistic factors—are as follows.

(i) The phonological matrix of the word form to undergo syllable elision must contain more than two syllables.

(ii) The syllable to be elided copies one or more distinctive components present in the preceding or in the following, or both, syllables. In extreme cases the repetition of constituents amounts to a full segment or even to the whole syllable. Notice that the mechanism of restructuring based on the phonotactic resemblance in the syllabic pattern, i.e. feature copying, has its parallel in ancient phonological processes, too, see lexicalized forms such as *pince* ⟵ *pivinica* 'cellar' or *pénz* ⟵ *p^{j}enez* 'money', and its pure form of realization is haplology, cf. *számkivet* ⟵ *számkive vet* 'banish' or *közvetlen* ⟵ *közvetetlen* 'indirect'. This criterion as the source of syllable elision in present-day Hungarian indicates that FSSE converges to haplology (see point **3**, below).

(iii) Syllables subject to FSSE cannot be in a phonotactically extreme position, that is, adjacent to active morpheme boundaries or in a peripheral position within the phonological matrix (cf. [va(:)l(:)ɑ:t] ⟵ *vállalat* 'industrial company'). On the other hand, the inactivation of morpheme boundaries induces syllable elision in casual speech such as in [meterolo(:)giɑ] ⟵ *meteorológia* 'meteorology', [bɑkterolo(:)giɑ] ⟵ *bakteriológia* 'bacteriology', and so on.

(iv) The tendency of syllable elision is incongruous with the highlighted (focus and/or stress) position of the syllable within the sequence.

(v) The restructuring through resyllabication of the word form as a whole proceeds by basically retaining the original architectonics of the input structure, cf., for example, *vállalat* ⟶ [va(:)l(:)ɑt], *vállalatot* ⟶ [va(:)l(:)ɑtot] 'industrial company-ACC', vs. *vállalatot* ⟶ *[va(:)l(:)ɑt] or *[va(:)l(:)ot]).

(vi) Syllable elision in casual speech often leaves behind phonetic traces, i.e. fragments of constituents of a segment or a reduced segment, cf. *társaság* ⟶ [ta:rʃ(:)a˙g ~ tarʃʃa(:)g ~ ta:ʃag] 'company'.

(vii) Types of item under point (vi) represent vacillating cases sensitive to tempo, prominence and extralinguistic circumstances.

(viii) The occurrences of syllable elision do not exhibit a systematic relationship with the phonetic quality features of the nucleus. (Recently, Siptár (1991, 38) gave the opinion that there would exist such a relation in the sense that the presence of a [+high] vowel in the nucleus makes the tendency more attractive. With Bárczi's description in mind, see above, this suggestion is not abhorrent to common sense. However, I do not find strong evidence to share Siptár's view.)

3. As is obvious from the lists of the main characteristics of TOSS and FSSE, respectively, the two process types represent different variants of contracting sequences within a pair of word boundaries. This is true even if they seem to be identical, especially with regard to the common result of the loss of syllable in the outcome. The main points of the differences may be summarized as follows.

(i) In its generalized form the occurrence of TOSS is conditioned upon the presence of two adjacent open syllables in the sequence (but cf. point **1**(iv) above, according to which some [+cont] consonants behave as if they did not make a syllable closed). It is blocked otherwise. In contradistinction to TOSS, the phonotactic attribute in question functions in FSSE as an indicator only, not as a criterion, cf. for example, [[ᵊzɑs:ẽⁿtet:e]] ⟵ *ez azt jelentette* 'this meant that'.

(ii) What really happens in TOSS is vowel elision whereas in FSSE we are faced with syllable elision proper. This claim needs, however, two remarks. There are instances in which FSSE also gives rise to vowel elision such as in [[ɑzta˙n]] (*sic*) ⟵ *azután* 'then' failing to reduce the rest of the sequence. On the other hand, in a number of instances an intermediary phase between vowel loss and syllable elision may also occur, see further below, which former may be held to be an anticipatory variant of the latter procedure.

(iii) The domain of FSSE is not confined to the deletion of vowels and not even to that of segments with a [+voice] component, cf., again, *társaság* ⟶ [tarʃa(ˑ/:)g] 'company'. This is in order to remark here that this type of deletion is conceived of by Kálmán (1988, 10) in another way. In his view what in fact happens is that between two [–voice] consonants every [+voice] constituent will be deleted from the sequence by FSSE and, simultaneously, the process results in a new, reintegrated syllabic pattern. It is, however, not the case for two reasons. First, not only [+voice] components but also [–voice] ones will be dropped, let alone the fact that some [+voice] constituents, more precisely the [NAS] dimension, will never disappear. Secondly, as a consequence of FSSE, the truncated sequence can exhibit what I called phonetic traces that considerably influence the syllabification of the remainder of the input sequence. What is labelled by Kálmán the reintegration process of the syllabic pattern is a metaphor for the mechanism of the organisation of sequential chains at a level higher than that of the syllable (also cf. point 4, below).

(iv) FSSE includes the possibility of deleting more than one syllable, see, e.g. *szoc*[V]*sta* ⟵ *szocialista* 'socialist', in which the most common realisation of [V] is [ɪ], that is, a reduced, somewhat centralized [+high +front –round] short vowel.

(v) FSSE, as a process, belongs to a quite different framework of causal relationships than does TOSS (see, again, points 2(ii) and 2(iii)).

4. In order to give a general evaluation of the mechanism of how FSSE functions I rely on the idea of what I called the 'Global Programming Principle' [GP] (cf. Szende 1992, 168–180). The crucial point is that in syllable deletion GP reduces the redundancy of the phonological matrix, i.e. the underlying input representation, of the word form by reducing the informational space of the articulation process which it manifests itself in. The way GP functions is to scale down the number of possible constituents actually available in speech production. As a consequence, the details of the articulation program will not completely cover the whole set of the constituents the input phonological representation is built up of, leaving in this way behind gaps in the implement. (With regard to the special requirements of the present issue some general claims of GP should be conjured up in this section.)

> The term 'global programming', in a strict sense, is a metaphorical expression of the idea that the components between extreme boundary markers of a word form (or a sequence of word forms, in a somewhat looser sense) of a phonological representation matrix are contained in a network

of mutual dependencies. A basic category and also the main concern in the GP principle is a specific interpretation of the phonological representation as a matrix which the implementation program is based upon. The main points:

(i) A phonological representation [PR] is a structure bounded by word boundary markers. This means that there are postlexical accommodation rules that, in normative lento speech production, apply word internally but not across a word boundary. (In Hungarian, for instance, *-t*, *-d*, *-n*, *-l* are commonly palatalized by a following *-j(-)* if what is between them is a morpheme boundary, but if the input segment is separated from the trigger by a word boundary, palatalization fails to apply in the first stage.)

(ii) The set of elements in a PR is an ordered set. This means that (ii/a) the number of elements permitted to occur between a pair of word boundaries is limited (with respect to Hungarian, cf. e.g. Szende 1976, 159); (ii/b) the choice of the order of elements in a PR is primarily determined by phonotactic constraints defined by syllable structure (segment sequence constraints for Hungarian are a rather unelaborated chapter of Hungarian phonology, though cf. Siptár 1980, Kassai 1981, Hell 1987, Törkenczy 1989); (ii/c) PR matrices are characterized by 'supersegmental' ordering constraints (in Hungarian the most widespread of these is vowel harmony).

(iii) Except for lento–normative utterances, the PR matrix (as a bounded structure of ordered elements) surfaces with structural distortions of definite types: distortions "in the first stage" apply within the bounds of a word-level PR matrix, keeping the original vocal pattern (called by Schnitzer (1972) the *Schallgestalt*). Lenition phenomena, but also data taken from first language acquisition processes and slips of the tongue offer further independent proofs for the fact that word forms are of a holistic nature. (Also, such holistic nature has been recently supported by an investigation exploring the time span needed for the access of a unitary 'image' (*le groupe rythmique comme signe structural de langue*) in what is called 'feed-forward', cf. Kojima (1991, esp. 333).) The global realisation program of a word form, then—again in the first stage—extends no further than the nearest word boundary.

(iv) The realisation program simultaneously excites all articulatory constituents involved in the word form for the whole time span of the sequence. In accordance with (iii), then, the planning of speech production takes place, in an elementary sense, in terms of an invariant word form—Hörmann's (1970/1971, 248 ff.) term is *Impulsfigur* or *Plan*, Linell's (1979, 48) is *(phonetic) plan*, Dressler's (1985 *passim*) corresponding term is *frame*. This view completely matches the biological notion with respect to the physiology of articulation that articulatory movements are synergistic, i.e. the production of speech units is based on pre-programmed muscular activity (cf. Craik 1947–1948 and Stetson's (1928) motor theory; the same has been supported by a number of more recent contributions, including Bergmann's (1987, esp. 106) theorem of word form internal time compression (*Isochronie-Tendenz*)).

(v) PR as used in natural speech production has been determined by three 'operational attributes' that interfere with the ideal form of a full PR matrix. (v/a) the columns of constituents defined by individual phonemes and collectively making up PR have been claimed to have 'permeable' dividing lines between them; (v/b) the relative constancy of forms has been eliminated to a large extent; and (v/c) the inventory of constituents has been significantly reduced. The last point includes the claim that the remaining components of phonetic identifying features do not necessarily fill in the temporal niche assigned to the corresponding phonological unit but may be transposed to those of other phonological units, leaving their own phonological niches empty. (For details and a more comprehensive argumentation, see Szende 1992, 168–180.)

The global programming of articulation processes resulting in the elimination of (some of the) subphonemic components and the partially unordered rendering of constituents provides us with the explanation of why and how transient forms within the occurrences of possible variants in the distorted word forms, with the same PR, emerge. Namely, with respect to individual realisations the program turns out not to be "global" in the same degree: at some points of the process it will show up as a fine-grained, high-resolution one whereas at some others it will be a rough, low-resolution one due to the actual strategic requirements of speech communication. This is, however, the area which the present comparison cannot be extended over. Unfortunately, TOSS and FSSE cannot be investigated on the basis of the same type of data sets. Taking all this into consideration, the bounds of the comparison may be summarized as in point **5**, below.

5. A comparison of TOSS with FSSE, two types of lenition processes, can only be carried out on the basis, and accordingly with the results, of assumptions of great probability value. Also, the expected consequences may only be hoped to have been a number of *per analogiam* type claims. Limitations in exclusively gaining solid statements lie in the fact that we are faced with (i) different types of data set, (ii) two nonhomogeneous variants of language with disparate morphonological structures, and (iii) necessarily diverse aspects of possible approximations of the individual phenomena selected for a comparative analysis.

6. Despite the limitations as in **5**(i–iii), the comparison provides several, and at least to some extent trustworthy, direct conclusions and, by implication, a few further consequences. Let us summarize them as follows.

(i) Syllable elision in FSSE ultimately results from the realisation of the primary difference existing between the intensity values of marginal/peripheral (and especially of the first, virtually stressed) syllables on the one hand and those of the word-internal (non-marginal) syllables, on the other. This is evidenced by the fact that the reduction of vowels and also that of adjacent consonants results eventually in the same effect making in this way reduction to be conceived of as a spring-board to syllable elision.

(ii) In Proto- and Old Hungarian, as was pointed out by Bárczi (1954/1958, see above), the tendency of 'two-open-syllable shortening' seems to have been one of the devices to implement the respective differences of intensity relationships within the sequence. This is not shown, nevertheless, by the mere fact of the occurrence of vowel elision—what would in itself only be quite a tautological, or at least a rather weak, argument in this respect—but by the fact that vowel elision may be mora preserving, i.e. it is often accompanied by compensatory lengthening (cf., again, the instances given in **1**(iii)).

(iii) Notwithstanding, TOSS in Proto- and Old Hungarian could only have emerged in connection with the demarcation of stressed positions, considering the following. (iii/a) The reduction of vowels, and the total loss of vowels later on, specifically in word final positions (such as in TihAl. (1055): *utu* 'road') caused by the lower action potential of articulation, hence the decrease of intensity in the acoustic output, cooccurs with the decrease of intensity of vowels in the first, i.e. stressed syllabic positions. Whenever this is the case, the distinction between stressed and unstressed vowels will remain upheld; (iii/b) The other side of the coin is that the presence of a reduced vowel in the ambient syllable enabling the vowel in the stressed syllable to uphold the same contrast with a lower degree of intensity, will enhance the tendency to lowering the action potential of articulation in the whole sequence.

7. For some further consequences it will be taken for granted that (i) changes within a pair of word boundary markers go back indeed to the source identified in (iii/a)—being supported by the circumstances mentioned in (iii/b), above—in accordance with what has been said under point **6**(ii). In addition, (ii) the role of the phonemic/phonotactic constituency and, at the same time, the demarcating function of word final reduced vowels get gradually lost (cf. the constraints in **1**(i)). Finally, (iii) TOSS ended up with a lower capacity in terms of intensity in view of the whole sequence to undergo the process (see point (iii) in this section). The premisses in 7(i–iii) offer the following inferences.

(i) The unified effect resulting from the loss of one of the original function of vowels becoming reduced (cf. point 7(ii)) and TOSS necessarily lead in the historical period of Old Hungarian to a much higher degree of the integrity of sequences in speech production. This tendency of change was complemented by the alteration of morphonological structure, too. Original bisyllabic substantives consisting of a root morpheme and a formative such as the predecessors of *-ból/-ből*, *-tól/-től*, etc.—in the process of integration on the way to become formatives—undergo a dramatic truncation via destressing (cf. E. Abaffy 1974, 439–440). The phonological–phonotactic pattern of these morphemes bears witness against their coming under the rule of TOSS, cf. (Proto-Hungarian) $rokV_{[\mathrm{back}]}$ + *-l* ⟶ *-ról/-ről*, *t(i/y)ġe* + *-l* ⟶ *-tól/-től*, $bœlV_{[\mathrm{front}]}$ + *-l* ⟶ *-ból/-ből*, etc. that underwent another mechanism of distortion, namely the same that turns present-day *vállalat* into [vaːlːɑt]. Notice that *-ból/-ből* originally shows the same phonotactic pattern with feature copying as does its present-day parallel *vállalat* and also, *-ról/-ről* or *-tól/-től* in their original versions contain a similar [+voice +pharyng/laryng] segment *H* as does present-day *tehát* and pass the same stages of phonotactic restructuring as *tehát* (⟶ [ta(ˑ/ː)t]) lenites (cf. Szende 1992, 208–209). Considering the same communicative, and in particular, semantic background of the change in this group of morphemes, the underlying principle of the mechanism—narrowing the informational space—should have been identical to that of present-day FSSE, too (cf. 4(i–iii) again). Accordingly, we could say, beside TOSS also FSSE in a qualified sense was at work in early historical times. (The qualification here means that content words were not involved.) On the other hand, TOSS-type instances of lenition can be observed in present-day Hungarian as well, although they are rare cases: some adverbs *(azután* 'then' ⟶ [ɑztaːn] (*sic*) or *természetesen* 'naturally' ⟶ [termeːsetʃen]) in unstressed positions can undergo vowel loss.

(ii) On grounds of the foregoing a second major consequence may be drawn sounding like this: a '(more) isolated ⟶ (more) fluent' type of change in the basic mechanism of speech production took place or, more precisely, was being accomplished in Old Hungarian. As is obvious without further proof, this basic alteration made room for a great variety of secondary changes in the suprasegmental structure of the language, the most important of which was the deliberation of the variability of phrase structure types and, as a consequence, the emergence of further, more fine-grained syntactic and semantic structures.

(iii) A secondary concomitant of the changes suggested herein should have been the increase of speech tempo in accordance with the lower quantity of the total output per sequential units to be produced. With this, again, a vicious circle will have been opened up: the increase of speech rate goes with the increase of the number of subsequent changes as was pointed out for Hungarian by Kubínyi (1958).

Finally, to answer the question in the title, particularly with regard to what has been said in points 7(i) and 7(iii), the two process types are in the relation of partial overlapping with an intersection such that set A of TOSS phenomena and set B of FSSE phenomena have some members in common: $A \cap B$ = formatives/form words.

References

E. Abaffy, E. 1974. Szóvégrendszerünk az ős- és az ómagyar korban [The system of word final units in Proto- and Old Hungarian]. In: Magyar Nyelv 70: 430–40.

Bárczi, G. 1958. Magyar hangtörténet [A historical phonology of Hungarian]. Akadémiai Kiadó, Budapest. (First edition: 1954).

Bergmann, G. 1987. Steuerungsprozesse des Sprechens. Helmut Buske Verlag, Hamburg.

Craik, K. 1947–1948. Theory of the human operator in control systems. In: British Journal of Psychology 38: 56–61, 142–48.

Dressler, W.U. 1985. Morphonology. The Dynamics of Derivation. Karoma Press, Ann Arbor.

Hell, Gy. 1983. A magyar tőszókincs fonémastatisztikája, szótagtípus statisztikája és mássalhangzó-kapcsolódásai [A statistical investigation of the occurrence of phonemes, syllable types and consonant clusters in Hungarian stems]. Ph.D. dissertation. Budapest.

Horger, A. 1911. Egy ismeretlen magyar hangtörvény [An unrevealed Hungarian sound law]. Nyelvészeti Füzetek 65. Budapest.

Hörmann, H. 1971. Psycholinguistics. Berlin–Heidelberg–New York. (Original edition: Psychologie der Sprache. 1970).

Kálmán, L. 1988. Derivational monotonicity and simplex features in phonology. Ms. Budapest.

Kassai, I. 1981. A magyar beszéd hangsorépítési szabályszerűségei [Phonotactic regularities in Hungarian speech]. In: Magyar Fonetikai Füzetek—Hungarian Papers in Phonetics 8: 63–86.

Kojima, K. 1991. Dans quelle mesure peut-on prévoir la séquence phonétique qui se réalise en image? La voix intérieure qui précède l'émission de voix. In: Actes du XIIème Congrès International des Sciences Phonétiques. Vol. 5, 330–3. Aix-en-Provence.

Kubínyi, L. 1958. Magyar nyelvtörténeti változások vélhető összefüggéséről [On possible interconnections of Hungarian sound changes]. In: Magyar Nyelv 54: 213–32.

Linell, P. 1979. Psychological Reality in Phonology. Cambridge University Press, Cambridge–New York–London–Melbourne.

Losonczi, Z. 1923–1927. A magyar tővéghangzók [The stem final vowels of Hungarian]. In: Nyelvtudományi Közlemények 46: 242–86, 402–58.

Melich, J. 1910. A tővégi magánhangzókról [On stem final vowels in Hungarian]. In: Magyar Nyelv 6: 150–59.

Pais, D. 1952. Imperativus a nyelvészetben [Maxims in the science of language]. In: Magyar Nyelv 48: 119–50.

Saussure, F. de. 1968. Cours de linguistique générale. Payot, Paris. (First edition: 1916).

Schnitzer, M. 1972. Generative Phonology—Evidence from Aphasia. University Park, Pennsylvania.

Siptár, P. 1980. A note on initial clusters in English and Hungarian. In: Acta Linguistica Hungarica 30: 327–43.

Siptár, P. 1991. Fast-speech processes in Hungarian. In: Gósy, M. (ed.): Temporal Factors in Speech, 27–61. MTA Nyelvtudományi Intézet, Budapest.

Stetson, R. 1928. Motor Phonetics. Amsterdam.

Szende, T. 1976. A beszédfolyamat alaptényezői [Fundamental factors in speech]. Akadémiai Kiadó, Budapest.

Szende, T. 1992. Phonological Representation and Lenition Processes. (Magyar Fonetikai Füzetek–Hungarian Papers in Phonetics 24.) MTA Nyelvtudományi Intézet, Budapest.

Törkenczy, M. 1989. Does the onset branch in Hungarian? In: Acta Linguistica Hungarica 39: 273–92.

Address of the author: Tamás Szende
Research Institute for Linguistics
Hungarian Academy of Sciences
Színház u. 5–7.
1250 Budapest, P.O. Box 19.
Hungary

Acta Linguistica Hungarica, Vol. 42 (1–2), pp. 75–82 (1994)

HAS SPEECH TEMPO SPEEDED UP IN HUNGARIAN OVER THE PAST 100–120 YEARS?

ILONA KASSAI

In the phonetics literature and professional discussions mention has often been made over the past three decades about the possible acceleration of speech tempo in standard Hungarian. This hypothesis is not at all unfounded as acceleration of speech tempo from period to period may provide sufficient explanation for several historical changes (cf. Kubínyi 1958; 1961).

To my knowledge, two attempts have been made to give an indirect proof of that hypothesis. Vértes O. (1987, 38; 1988a; 1988b; 1989, 382), in support of the tendency of acceleration, referred to rate values published in the literature and calculated for speeches of parliamentary sessions in various years as recorded in shorthand writing. These values were complemented by Szende (1987, 46) with data from 1986, with reference to the Hungarian life champion of shorthand writing. All these data are summarized in Table 1 where, beside the values given in syllable per minute, I calculated the corresponding values in syllable per second.

Table 1

Year	Rate of speech	
	syllable/minute	syllable/sec
1869	120–140	2.0–2.3
1889	220	3.6
1924	240–250	4.0
1986	250–300	4.2–5.0

The above figures, though they reflect the tendency of acceleration in a seemingly convincing way, are not completely reliable. The reasons are as follows. First, the unit of speech rate used by stenographers, the syllable, only gives rough approximations about speech tempo as the syllables of Hungarian exhibit great structural variety. This means that, e.g., a six-segment syllable is hardly comparable with a two-segment syllable. On the other hand, a shorthand record does not contain any information about the actual realization

1216–8076/94/$ 4.00

of phonological syllables as stenographers transform their shorthand symbols into orthographic forms.[1] This fact places another restriction on using the syllable as a unit of tempo measurement. Namely, shorthand writing does not reflect eventual pauses and this strongly affects speech rate as tempo perception is highly dependent on the number and duration of intervening pauses.[2] For illustration I refer to the values calculated by Hegedűs (1957) for newsreading and sports broadcasts (Table 2). Therefore, it may easily happen that one Member of Parliament pauses less and produces more syllables during his speech than another member.

Table 2

	AT	sound/s	PT	$P_{average}$
newsreading	65 970	12	24 800	708
sports broadcasts	127 260	14	15 860	323

(AT = articulation time in ms, PT = pause time in ms, sound/s = speech rate, $P_{average}$ = average pausing in ms).

As can be read off Table 2, while in newsreading the ratio of pause time to speech time is only 1:2.4, in sports broadcasts it is 1:8.

The other attempt to prove the hypothesis of tempo acceleration has been made by Mária Gósy (1988)[3] who tried to determine the relationship between tempo perception and speech comprehension. In her first experiment she investigated the effect of articulation rate and fundamental frequency modulation on the perception of tempo by means of various versions of a synthetized basic sentence. This basic sentence was generated in five different versions in terms of F_0 modulations and the five versions were speeded up to 1.2 times and to 1.6 times the original tempo. One of the conclusions of the study is that "the acceleration of Hungarian speech is demonstrable at least indirectly, at the level of perception" (103). The conclusion is based on the behaviour of the

[1] For the interrelations of oral language and shorthand writing see Nosz (1963) and Szépe (1963).

[2] The notion of rate can be meant as: (1) articulation rate, i. e. the ratio between the amount of time devoted to the production of a given sequence of speech sounds and the number of speech sounds or syllables produced and (2) speaking rate, i. e. the number of speech events (sounds or syllables) is given as a function of the total speaking time comprising both articulation and pausing (cf. Kelly 1949).

[3] Throughout this paper I refer to the original Hungarian version as it was the only available source at the time of writing. Since then the investigation has been summed up in English (1992, esp. 169–79).

listeners involved who judged 'normal tempo' the variant which displayed an acceleration value of 1.2 as opposed to the basic variant. On closer inspection, however, one can see in the data (*op. cit.* 102) that the rate given for the basic sentence is 9.2 sounds/s which characterized the rate calculated by Fónagy–Magdics in 1960 (p. 186) for reading poems (Table 3). The rate of the first speeded up version is 11 sounds/s and this corresponds to the average standard found by Fónagy–Magdics (*op. cit.* 185).

Table 3

Speech styles	Average speech rate sounds/s
Reading poems	9.4
Tales	9.57
Prose	10.73
Children's talk	11.68
Conversation	12.89
Sports broadcasts	13.83
Average	11.35

Thus, it is not surprising at all that it was this sentence that listeners judged normal for its tempo. Further, as the second version with a tempo speeded up by 1.6 was judged fast, we can state that 15 sounds/s characteristic of this sentence variant is considered as fast rate in everyday usage. Therefore, this experiment, if we interpret its data appropriately, does not help us answer the question asked in the title of this paper.

A viable and reliable piece of indirect evidence may be drawn from the famous experiment conducted by Meyer–Gombocz in 1909. They measured the duration of individual segments as a function of the length of utterance in 5 Hungarian words derived from the same verb root. The experiment was replicated by Tarnóczy in 1965.[4] These two attempts seem to be relevant for our purposes as (1) the influence of pauses is excluded, (2) the language sample is identical, (3) the speaker is, in both experiments, of about the same age, education and socioeconomic status, (4) the laboratory circumstances of the experiment do not allow for reduced articulation and (5) the two experiments embrace about 60 years, i. e. two generations' time. Thus, the conditions by which Szende (1987, 19) considers acceleration as documentable can be ful-

[4] His aim was to define the role of segmental durations in automatic speech recognition (1965).

filled at least in part: "A near authentic documentation is only possible if we gain calculation or estimation based on a large spoken corpus guaranteeing statistically evaluated results referring to a period of time. A further condition is that the corpus must be recorded on the same or very similar topic, in identical speech situations." It follows that the superposition of the tables reflecting the two experiments promises important conclusions (Table 4).

Table 4

Words	Duration of vowels (ms)									
	a:		o		ɔ		o:		ɔ	
	1909	1965	1909	1965	1909	1965	1909	1965	1909	1965
tát	272	210								
tátog	242	180	118	155						
tátogat	209	140	94	95	109	115				
tátogatók	190	120	79	85	87	105	221	130		
tátogatóknak	182	110	79	80	86	90	172	110	–	140

Looking at columns we can state that the longer the sequence, the shorter the duration of individual segments. Horizontally it becomes clear that the farther the segment is from the beginning of the sequence, the longer its duration. The values of [ɔ] vowels in the 3rd and 5th syllables are convincing in this respect: in the 5th syllable the [ɔ] is 55% longer than the same vowel in the 3rd syllable. What is more, it is 27% longer than the phonologically long [a:] in the 1st syllable. The comparison of the values obtained from the two experiments yields the conclusion that, since the beginning of the century, the durations of short and long vowels have got closer to each other to a considerable extent. Duration ratios calculated by Gombocz in 1909 for the [o/o:] pair in *tátogatók* and *tátogatóknak* are 1:2.79 and 1:2.17, respecively. The same ratios turn out to be, 60 years later, 1:1.52 and 1:1.37. It is worth noting that in the 2nd and 3rd syllables the values of the later experiment exceed those of the former one.[5] This fact suggests that durational values have got closer not only by the shortening of long vowels but also by the lengthening of short vowels, short-

[5] The heuristic value of this result is inceased by the fact that Gombocz, unlike Tarnóczi, segmented the sequence in an unusual manner: he considered the implosive phase of stop consonants as belonging to the preceding vowel while the explosive phase to the following vowel (see Meyer–Gombocz 1909, 5). In other words, the values given for vowels by Gombocz are more telling than those given by Tarnóczy.

ening being the stronger tendency.[6] To further interpret the data we make use of the observations and research results by Klára Magdics (1969) who observed, in connection with the influence of tempo on segmental duration, that the longer the duration of vowels with normal tempo the more shortening is found with fast tempo and conversely. In view of these results the marked shortening exemplified from the beginning of the century up to 1965 in the integrated table can only be explained by the acceleration of the tempo of Hungarian speech. The reasoning may be confirmed by referring to another fact. In Lannion, France, experts of speech synthesis reported that in longer sequences, toward the end of the sequence they perceive a gradual speeding up in tempo. A plausible explanation is the following. The text-to-speech synthesis system of CNET[7] uses, as building blocks, diphones whose parameters, duration among them, can be automatically varied between fixed minimum and maximum values, the actual value being the function of the position within the sequence (cf. Stella 1985, 438–50; Bartkova–Sorin 1987; Bartkova 1990). If so, durational variation accompanying the increase of utterance length in natural speech meets restrictions in synthesis and provokes, thereby, a sensation of acceleration. That is to say, tempo acceleration is related to shorter-than-required duration.

Having arrived at this point in seeking an answer to the question of the title, it is quite natural that one is eager to know what the experiment initiated by Gombocz and Meyer in 1909 looks like three generations' time later, in 1992. Thus, I repeated the experiment, keeping the requirement of the sameness of conditions in mind. I then added the processed data to those of the two former experiments (Table 5). The 1992 data seem to corroborate the relationships revealed in the two former cases and give a direct illustration of the further shortening of the duration values. Moreover, if we calculate duration ratios for the vowels of adjacent syllables (Table 6), we can document a narrowing time span between short and long vowels manifesting their convergence to a central value. This levelling of the values decreases diversity in timing and the resultant monotony leads to an impression of acceleration in tempo. The impression itself can be explained by the nature of perceptual processing which gets speeded up when working with a more homogeneous speech signal.

[6] I could reveal the same tendency in my investigations concerning timing phenomena of Hungarian speech: lengthening effect of the factors conditioning duration was stronger for short vowels while shortening effect was stronger for long vowels (Kassai 1979, esp. 29–36).

[7] Centre National d'Etudes de Télécommunication.

Table 5

Interrelations of segmental duration and utterance length in 1909, 1965 and 1992

Words	Duration of vowels (ms)														
	a:			o			ɔ			o:			ɔ		
	1909	1965	1992	1909	1965	1992	1909	1965	1992	1909	1965	1992	1909	1965	1992
tát	272	210	227												
tátog	242	180	176	118	155	10									
tátogat	209	140	141	94	95	78	109	115	102						
tátogatók	190	120	117	79	85	58	87	105	94	221	130	117			
tátogatóknak	182	110	117	79	80	55	86	90	78	172	110	102	–	140	94

Table 6

Duration ratio of vowels in adjacent syllables (%)

Words	Duration ratio of vowels (%)											
	a:/o			o/ɔ			ɔ/o:			o:/ɔ		
	1909	1965	1992	1909	1965	1992	1909	1965	1992	1909	1965	1992
tátog	0.49	0.86	0.63									
tátogat	0.45	0.68	0.55	1.16	1.21	1.31						
tátogatók	0.42	0.71	0.50	1.10	1.24	1.62	2.54	1.24	1.24			
tátogatóknak	0.43	0.73	0.47	1.09	1.13	1.42	2.00	1.22	1.31	–	1.27	0.92

In conclusion, on the basis of the three experiments, it can be stated that the tempo of Hungarian speech has really accelerated, at least compared to the beginning of this century. The sensation of acceleration is basically due to two factors. On a more elementary level, the decrease of absolute durations is perceived as acceleration. Indirectly, monotony caused by the modifications of internal timing relations is felt as tempo acceleration. If we want to know what the overall speech rate may be in present-day Hungarian, one possible value comes out of the second experiment conducted by Mária Gósy in her study referred to above. The natural speech sample involved in her listening expermiments has yielded a value of 14 sounds/s to represent normal tempo (Table 7). This rate, by and large, corresponds to that calculated 30 years ago by Fónagy–Magdics (*op. cit.* 186) and Hegedűs (*op. cit.* 18) for sports broadcasts.

Table 7

Perceived tempo	Speech rate		Change of rate (%) $\frac{\Delta T}{T}$
	sound/s	syllable/s	
normal	13.98	5.97	-
fast	23.68	10.16	59
very fast	29.52	12.69	47
very slow	7.2	3	194

References

Bartkova, K.-Sorin, Ch. 1987. A model of segmental duration for speech synthesis in French. In: Speech Communication 6: 245-60.

Bartkova, K. 1990. Règles du modèle prédictif de la durée segmentale. Document de travail 1095/LAA/TSS/RCP. Manuscript.

Fónagy, I.-Magdics, K. 1960. Speed of utterance in phrases of different lengths. In: Language and Speech 3: 179-92.

Gósy, M. 1988. Tempóészlelés és beszédmegértés [Tempo perception and speech understanding]. In: Műhelymunkák a nyelvészet és társtudományai köréből 5: 87-122.

Gósy, M. 1992. Speech Perception. Hector, Frankfurt am Main.

Hegedűs, L. 1957. Sprechtempoanalysen im Ungarischen. In: Zeitschrift für Phonetik 10: 8-20.

Kassai, I. 1979. Időtartam és kvantitás a magyar nyelvben [Duration and quantity in present-day Hungarian]. Nyelvtudományi Értekezések 102. Akadémiai Kiadó, Budapest.

Kelly, W.J. 1949. Revised concept of rate. In: Journal of Speech and Hearing Disorders 14: 222-6.

Kubínyi, L. 1958. Magyar nyelvtörténeti változások vélhető összefüggéséről [On possible interrelations of historical changes in Hungarian]. In: Magyar Nyelv 54: 213-32.

Kubínyi, L. 1961. A beszédiram magyar nyelvtörténeti jelentőségéről [On the importance of speech tempo in explaining historical changes in Hungarian]. In: Magyar Nyelv 57: 337-40.

Magdics, K. 1969. A magyar beszédhangok időtartama nyugodt és gyors beszédben [Duration of Hungarian speech sounds at moderate and fast tempo]. In: Nyelvtudományi Értekezések 67: 45-63. Akadémiai Kiadó, Budapest.

Meyer, E.A.-Gombocz, Z. 1909. Zur Phonetik der ungarischen Sprache. Berling, Uppsala.

Nosz, Gy. 1963. A gyorsíró és az élőszó [Stenographer and oral language]. In: Magyar Nyelvőr 87: 96-104.

Stella, M. 1985. Speech synthesis. In: Fallside, F.-Woods, W.A. (eds): Computer Speech Processing, 421-60. Prentice Hall International, Cambridge, MA.

Szende, T. 1987. Megérthetjük-e egymást? [Can we understand each other]? Gondolat, Budapest.

Szépe, Gy. 1963. Nyelvészeti jegyzetek a beszédgyorsírásról [Linguistic notes on the shorthand recording of speech]. In: Magyar Nyelvőr 87: 231–7.

Tarnóczy, T. 1965. Can the problem of automatic speech recognition be solved by analysis alone? In: Rapport du 5e Congrès International d'Acoustique II. Conférences générales, 371–87. Liège.

Vértes O., A. 1987. Érzelmi világunk és a nyelv történeti változásának kölcsönhatása [Interactive changes in affect and language]. In: A Magyar Nyelvtudományi Társaság Kiadványai 179. Budapest.

Vértes O., A. 1988a. Interactive changes in affect and language: Some observations in Hungarian. In: Ural-Altaische Jahrbücher/Ural-Altaic Yearbook 60: 183–8.

Vértes O., A. 1988b. Időtényezők stílusbeli szerepe [Stylistic role of timing factors]. In: Műhelymunkák a nyelvészet és társtudományai köréből 5: 123–47.

Vértes O., A. 1989. Gedanken über die geschichtliche Veränderung des Sprechtempos. In: Proceedings of the Speech Research '89 International Conference. Magyar Fonetikai Füzetek/Hungarian Papers in Phonetics 21: 381–4.

Address of the author: Ilona Kassai
Research Institute for Linguistics
Hungarian Academy of Sciences
Színház u. 5–7.
1250 Budapest, P.O. Box 19.
Hungary

Acta Linguistica Hungarica, Vol. 42 (1–2), pp. 83–92 (1994)

VOICE QUALITY CHANGES IN OLD AGE

BOGLÁRKA BALÁZS

Introduction

To extend our life span is one of the oldest dreams of mankind. The prevention of diseases, the early detection of their symptoms, as well as the ever-increasing standards of therapy, keep raising the upper limit of human life expectancy. An increasing number of disciplines study the phenomena of ageing and the problems of old people. One of the oldest and most developed of these is gerontology or 'the science of ageing' that studies old age in terms of all relevant biological, psychological, and social aspects. Indeed, ageing is the totality and mutual interaction of all chemical, physicochemical, and morphological changes that a human organism undergoes during its lifetime and due to which it is less and less able to keep up its physiological equilibrium at a constant level. Gerontology, as part of physiology, used to study morphological and functional changes of the whole body. More recently its focus of attention has shifted to special ageing phenomena of the individual organs. The earliest examples of that specialization are the study of presbiopia—old people's failing eyesight—and presbiacusis—their defective hearing. The WHO defines 'old age' as starting at 65; 18.5% of the Hungarian population are over sixty-five years of age.

The most up-to-date definition of ageing has been suggested by Max Bürger: "Ageing is a biological process. It is an irreversible process of life that has its particular causes and mechanisms" (cf. Spellenberg 1985). The order of stages in biomorphosis is unalterable and no-one has solved the problem of rejuvenating an aged organism. Therefore, the most important objective of gerontological research and geriatric clinical practice is to make the period of old age as long and the old person as healthy as possible. Our organs of speech grow old as we do; it is of some importance, therefore, for us to get to know the phonation characteristics of healthy old people and the changes that their voice quality undergoes.

1216–8076/94/$ 4.00

The rudiments of phoniatry

An indispensable prerequisite of the complex process of phonation is neural control, embodied in a precise coordination of the muscles of speech organs (Schulz-Coulon 1980).

The basic factors of human voice production are

(i) pulmonic (egressive) airstream mechanism,
(ii) phonatory position, muscular tension, vocal cord vibration, and
(iii) vocal tract configurations acting as resonators.

Voice quality is determined by constitutional, hormonal, and psychic factors.

Human voice production is less than fully understood even today. Two competing views are widely held:

(i) the older but currently more influential aerodynamic-muscular theory and
(ii) the neurochronaxic or neuromuscular theory.

The aerodynamic-muscular view claims that vocal cord vibration is self-excited. Vocal cords in a phonatory position are tense. The pressure of pulmonic airstream will periodically open the glottis and the tense vocal cords keep jumping back to closed position again. The necessary vocal cord setting is neurally controlled.

The larynx participates in voice production by its musculature, by letting the vocal cords vibrate, and by fine movements of the cartilages themselves. Internal laryngeal muscles keep the vocal cords in motion and/or establish their phonatory position. The air expelled from the lungs sets the vocal cords in motion and voice is thereby produced. The direction of vibration is mainly horizontal but it has a less pronounced vertical component as well.

The sound wave produced by vocal cord vibration is transmitted to sublaryngeal and supralaryngeal cavities and forces its vibration on the air contained there (Frint–Surján 1982). The following parts of the vocal tract contribute to subsequent modification of the output: the trachea and the bronchi, the ventricles of Morgagni, the pharynx, the tongue, the velum, the nasal cavity and the accessory nasal sinuses. The vocal tract functions as a set of resonators but the soft parts of its walls also damp the voice a little (cf. Fig. 1).

Hoarseness is a voice production problem involving some additional noise components superimposed on the quasi-periodical voice produced by normal vocal cord vibration. Those noise components are due to imperfect periodicity of voice production. Two major factors that may result in non-harmonic waveforms are irregularities of vocal cord vibration and the turbulence of the airstream.

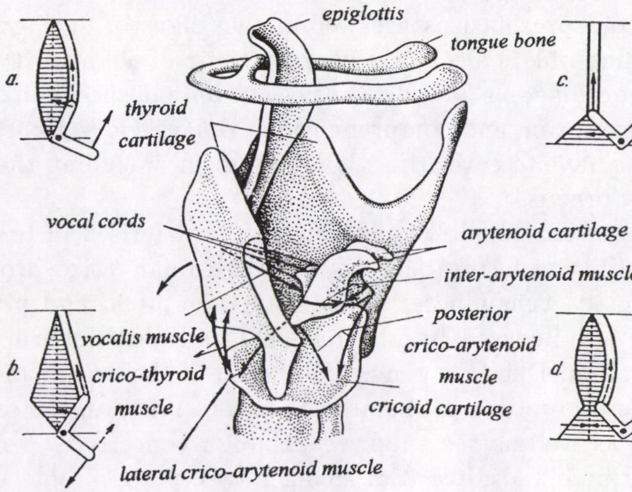


Fig. 1
The movements of vocal cords and laryngeal cartilages

In addition to the usual otolaryngological examinations, hoarseness can be explored by stroboscopy, glottography, and spectrography. The dynamics of voice production is examined by phonetography.

Phonation in various ages

The working of the larynx, hence human voice production, is under a hormonal influence during our whole life. Larynxes of new-born babies, whether they are boys or girls, are the same size. Small children's pitch, tamber, range of voice, and average tessitura are essentially uniform. This is known as 'child voice'.

The first change occurs during puberty, with hormonal maturation: voice-breaking introduces a difference between male and female voices (Schilling–Karthaus 1960/61). Men's voice goes an octave deeper and assumes a masculine tamber: it becomes characteristically different from women's. According to Gerhardt Böhme and Gudrun Minor's investigations (cf. Böhme–Hecker 1976) this so-called 'adult voice' can be heard between puberty and age 60. After that, the difference between the voice qualities of the two sexes gradually decreases: men's and women's tamber and pitch range, and even their singing voice become more and more alike. The height of female voice is reduced, the characteristic 'feminine' tamber wears off. The brazen resonance of male voice fades, masculine depth gradually vanishes, average pitch goes higher. Old people's voice is called 'old voice'.

Presbiopia and presbiacusis are well-known signs of ageing; but presbiphonia or 'sounding old' is also quite characteristic of old age. It is also to be kept in mind that old people have difficulties not only in speech production but also in speech perception and comprehension (Balázs–Gósy 1990). Therefore, our investigations have to cover the whole organism, including the interaction of its constituent organs.

The changes of **old people's voice** can be characterized in terms of quantitative and qualitative factors (Balázs 1992). Human voice production, including the intensity, tunefulness, tamber, average pitch, and pitch range of a person's voice is influenced by physiological and pathological processes of the hormonal system. Due to a general detrition of the organism, natural arteriosclerosis, the diminution of the active mass of muscles, the fatigue of the nervous system, as well as the subdued—though concerted—working of the system of endocrine glands, essential changes take place in old people's voice production.

What are those changes?

As the respiratory system grows old, lung capacity diminishes, the vigour of the heart declines. Thus, the 'pump' losing strength and the aerodynamic factor becoming less powerful, **voice intensity** decreases. In addition, the increasing sluggishness of laryngeal muscles causes subglottal pressure to diminish. Also, **sound sustension** capacity will reduce in length. Therefore, an old person's voice is substantially softer and he/she has to take a breath more often. Old people's voice production is characterized by some degree of feebleness, resulting in a trembling voice quality. As the external laryngeal muscles get worn out and the vocal cords lose some of their elasticity, average **pitch** and **range of voice** are also affected. As the tissues gradually calcify, the stroboscopic data are deformed. Reduced elasticity of vocal cords makes the range of voice narrow down. With respect to men this affects the low end of the pitch range: depending on the original register, the notes F, G, A, H tend to be lost. In the case of women it is the high end of the pitch range that is affected: the notes c2, d2, e2 are usually lost. Therefore, old peolpe's voice range is reduced to slightly more than a single octave, as opposed to the usual two octaves of adult voice: this makes old voice similar in this respect to that of children. Men's average speaking voice rises by $1–1^1/_2$ notes, whereas that of women either stays unchanged or lowers by half a note.

This is where smoking comes into the picture. Over the years, smoking causes some degree of thickening of the vocal cords; this cannot be taken to be a pathological deformity but the increased mass of vocal cords will make the voice go deeper.

Changes in the vocal tract

As the sound wave produced by vocal cord vibration proceeds along the vocal tract, it is transformed into the individual's characteristic voice by forcing its own vibration onto the resonant cavities above and below the glottis. Thus, the changes that the vocal tract undergoes as the speaker grows old will induce changes in his or her voice quality.

The calcification of the muscles and cartilages of the **trachea and bronchi** changes the quality and quantity of chest resonance.

The declining elasticity of the **ventricles of Morgagni** vitiates the acoustic parameters of the speaker's voice.

Swallowing difficulties and a partial opening of the **piriform recesses** make the tamber more 'veiled'.

The fatigue of **pharyngeal and laryngeal muscles** debases resonance and makes phonatory position harder to maintain.

The increasing difficulty of **tongue** movements makes articulation less precise. Also, it results in problems with the adjustment of the position of the larynx (its raising for high-pitched sounds and its lowering for low-pitched ones).

The movements of the **velum** are increasingly sluggish, whereby old people's speech is usually made more nasalized.

The mucous membranes of the **nasal cavity** show a tendency towards atrophy, another factor contributing to resonance changes.

Changes of voice quality

Physiological changes of the vocal tract in old age all add up to decrease the sonority of the speaker's voice. Acoustic analysis shows that its spectral properties undergo substantial changes with ageing. Sedlačkova (1966) and Balázs–Gósy (1992) present data suggesting that the average fundamental frequency of old voice differs from that of adult voice and that in old speakers' spectrograms formant bands tend to exhibit fuzzy—as opposed to clear-cut—edges.

The number of harmonics is smaller and the spectrum extends much less towards high frequencies. Formant areas of vowels contain some noise components. The older the speaker, the more it is the case that the harmonic components of front vowels, especially of [e] and [i], are broken up by noise components. (The reason for that is claimed by gerontologists to be presbiacusis, i.e. hearing loss of high-pitched sounds, as well as qualitative changes of hearing and imperfect acoustic feedback.)

Glottograms reveal jerky changes in F_0 intensity; the amplitude curve shows shorter intervals and the pattern of the waveform suggests muscular fatigue, cf. Childers *et al.* (1984) (Fig. 2).

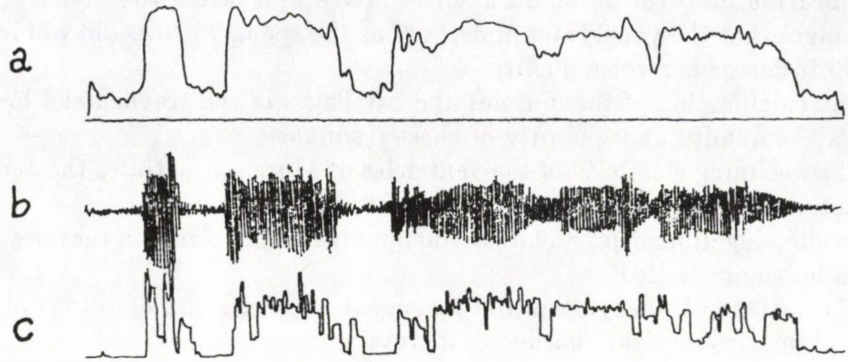


Fig. 2
Sample records of the speech of a 72-year-old woman
(a = intensity, b = glottogram, c = F_0 contour)

The dynamics of voice production is also subdued. Wendler's investigations (cf. Wendler–Seidner 1987) show that the shape of phonetograms of old speakers curiously resembles those of teenage subjects exhibiting sound-breaking: both pitch range and voice intensity are reduced (Fig. 3).

The highly complex process of voice production—involving speech-specific breathing, vocal cord vibration, and proper functioning of the vocal tract—is controlled and coordinated by the central nervous system on the one hand, and by the peripheral and reflex systems on the other. With growing age, degenerative cerebral changes and the wearing out of the peripheral nervous system may result in comprehension problems, too. The increased length of latency entails a qualitative impairment of voice production and of speech in general. Old people's speech is therefore less dynamic, duller, and more monotonous. Emotions play a lesser role in their voice production, although this may show highly individual variation in some cases. Their speech keeps dropping in volume and is rather fragmentary. The explanation Wendler offers is that the sensory, extrapyramidal, and limbic systems tend to be less efficient in controlling voice production.

A series of experiments has been conducted (Balázs–Gósy 1990) concerning speech comprehension in old age, to see if the general belief that old persons' comprehension difficulties are due to presbiacusis and/or necessar-

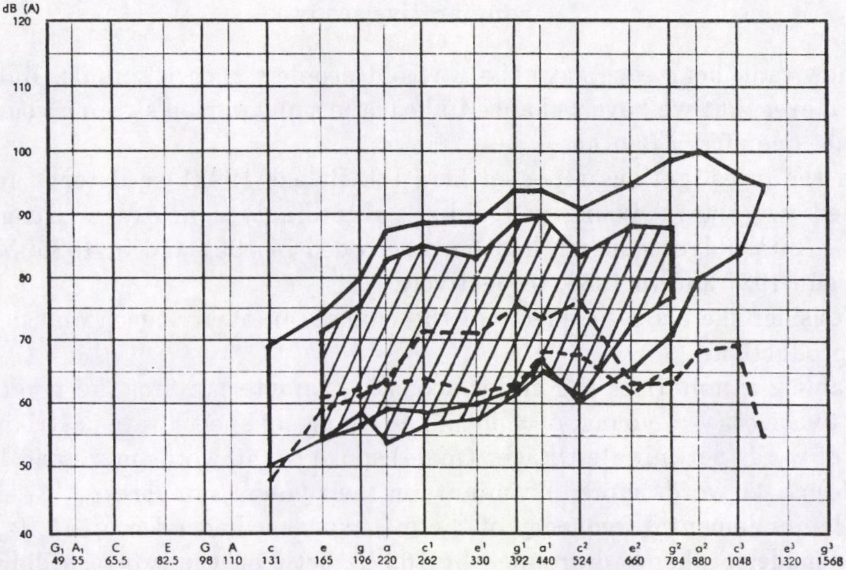


Fig. 3
Phonetogram of an old speaker (shaded area) compared to that of average normal speech (dotted lines indicate average pitch)

ily cooccur with hearing defects is correct. We asked twenty subjects between 70 and 80 years of age to listen to a short text taken from a tale and asked five comprehension questions. The ratios of correct answers are summarized in Table 1. Our data show that speech comprehension difficulties are rather grave in old age (obviously, the number of correct answers varies with the subject's age, actual hearing capacity, intelligence, and logical skills as well.) Where the control group (average age: 25) gave 90–100% correct responses, the group of over-70s reached an average score of 51% (Table 1).

Table 1

Speech comprehension in old age	
Question	Correct answers (%)
when?	16.6
why?	91.6
where?	87.5
how much?	33.3
for what reason?	25.0

A comparative study

We know (and hear each day) the way old people's speech sounds. But it is rarely if ever that we have a chance to hear the same person's young voice and old voice one after the other.

In the investigations reported here (cf. Balázs 1992) we listened to and analysed two actors' young and old-age recordings (courtesy of Hungarian Radio). Antal Páger's speech had been recorded in 1957 and in 1985, Margit Dajka's in 1937 and in 1983, respectively.

Consider the acoustic effects of the changes in their speech and in their voice production.

Table 2 summarizes the major acoustic characteristics of the two actors in the two ages (y = young, o = old). The change of speech rate is striking: in terms of words per minutes, Páger (male) spoke ca. 40 words more and Dajka (female) ca. 20 words more in a minute in their young samples.

The fundamental frequency of the male speaker had risen by 15 Hz, that of the female speaker had dropped by 160 Hz between the two recordings.

It is very characteristic that the majority of speech sounds fell within a frequency range 500 Hz narrower for Páger and 1,500 Hz narrower for Dajka than in their young speech.

Also, the frequency of pausing changed in a revealing manner: the male speaker used eight times, and the female speaker five times, as many pauses in old age as in their youth.

Table 2

Features	Age	Páger	Dajka
speech rate	y	12.3	13
(sound/s)	o	8.6	11
fundamental	y	90	280
frequency (Hz)	o	105	120
frequency	y	2,500	3,000
range (Hz)	o	2,000	1,500
number of	y	n	m
pauses	o	8n	5m

The acoustic characteristics of old vs. young speech can be summarized schematically as in Fig. 4. The upper figure suggests vigorous speech sound components, well-contoured formants and consonantal noise components that are easy to discern and are dispersed throughout the frequency band shown. The lower figure suggests diminished formant intensities, noise components through the whole waveform, vague formants, and blurred contours.

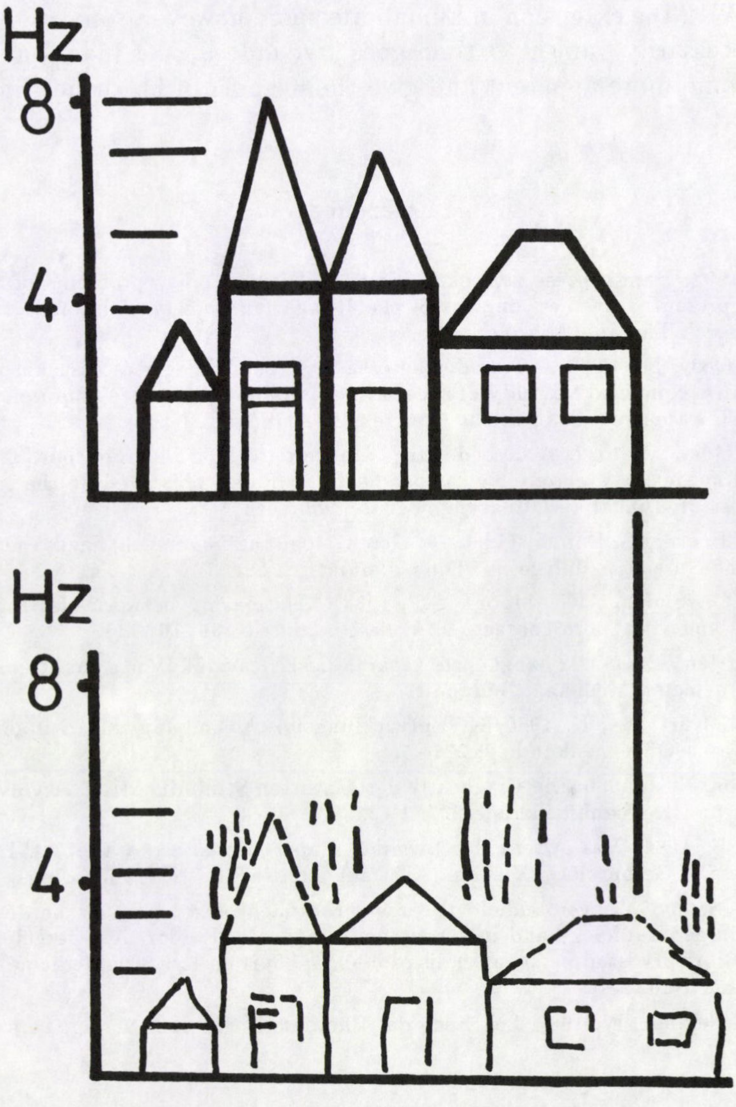


Fig. 4
A schematic rendering of acoustic features of young vs. old speech

Conclusion

As the organism grows old, it is not only hearing and sight that deteriorate but one's speaking voice, too, becomes veiled, colourless, trembling, and exhausted. With the extension of human life span, however, some of our abilities do not necessarily vanish as time goes by; indeed, the individual's virtues might become more apparent and give the evening of his or her life a special quality.

References

Balázs, B. 1992. Hangképzési zavarok időskorban [Voice quality problems of the elderly]. Paper presented at the Congress of the Hungarian Society of Phonetics, Phoniatry, and Speech Therapy, Miskolc.

Balázs, B.–Gósy, M. 1990. A beszédérthetőség és a beszédmegértés összehasonlítása idős korban [A comparative study of speech intelligibility and speech comprehension at an old age]. Paper presented at the Congress of Audiology, Nagykanizsa.

Balázs, B.–Gósy, M. 1992. A beszéd vizsgálata gerontológiai szempontból [The examination of speech in a gerontological perspective]. Paper presented at the International Congress of Hungarian Otolaryngologists, Debrecen.

Böhme, G.–Hecker, G. [Minor, G.] 1976. Gerontologische Untersuchungen über Stimmumfang und Sprechstimmlage. In: Folia Phoniatrica 22: 176–84.

Childers, D.G.–Smith, A.G.–Moore, G.P. 1984. Relationship between electroglottograph, speech, and vocal cord contact. In: Folia Phoniatrica 36: 76–118.

Frint, T.–Surján, L. 1982. A hangképzés zavarai, beszédzavarok [Voice quality problems and speech defects]. Medicina, Budapest.

Schilling, A.–Karthaus, K. 1960/61. Entwicklungsbeschleunigung und Stimmwechsel. In: Hals-Nasen-Ohrenheilkunde 9: 271–5.

Schulz-Coulon, H.-J. 1980. Die Diagnostik der gestörten Stimmfunktion. Archiv für Ohren-Nasen- und Kehlkopfheilkunde 227: 1–2.

Sedlačkova, E. 1966. Das Altern der Stimme. Paper presented at the 7th International Congress of Gerontology, Vienna. [Abstract published in the Proceedings, Vol. 4: 469.]

Spellenberg, S. 1985. A geroaudiológia és a gerootoneurológia aktuális kérdései [Current issues in geroaudiology and in gerootoneurology]. In: Káldor, A.–Gergely, F.–Kulka, F. (eds): Az orvostudomány aktuális problémái [Current Issues in Medicine], 21: 35–36. Medicina, Budapest.

Wendler, J.–Seidner, W. 1987. Lehrbuch der Phoniatrie. Springer Verlag, Leipzig.

Address of the author: Boglárka Balázs
'Szent János' Hospital
Diósárok u. 1.
1125 Budapest
Hungary

Acta Linguistica Hungarica, Vol. 42 (1–2), pp. 93–101 (1994)

A SPEECH IMPROVEMENT TECHNIQUE BASED ON VISUAL FEEDBACK*

KLÁRA VICSI–EMŐKE KOVÁCS-VASS–PÉTER BARCZIKAY

1. Introduction

In traditional speech therapy a process oriented approach is generally used (Luchsinger–Arnold 1965); the speech therapist gives instructions on how to use the speech organs while forming sounds. During normal speech development, children never receive instructions on how to move or where to place their speech organs (Povel 1991).

Instead of the process oriented approach, or to supplement it, we would like to offer a product oriented method where correct visual feedback is the most important cue (see Fig. 1). Nevertheless, during the sessions we use the patients' incorrect auditory feedback as well. The patients can see and hear the speech pattern at the same time. This system helps the patients to discover how to move their speech organs by simultaneously comparing the visual patterns (speech pictures) of the normal acoustic speech signal with the defective one.

2. General method of speech therapy

The aim of our system is the correction of defective dimensions of speech. The system evaluates those parameters given by the auditory model (Vicsi 1982; Zwicker 1982) which are important in speech quality: loudness changing in time, cochleogram (similar to the perceived one) changing in time, speech tempo differences, spectral maximum extractions, voiced-unvoiced decision, etc. (Vicsi–Tihanyi 1987; Vicsi *et al.* 1990). Patients may simultaneously modify and correct their defective articulation by real time interaction between speech sounds and computer displays.

* This research was supported by a Soros Foundation grant.

1216–8076/94/$ 4.00

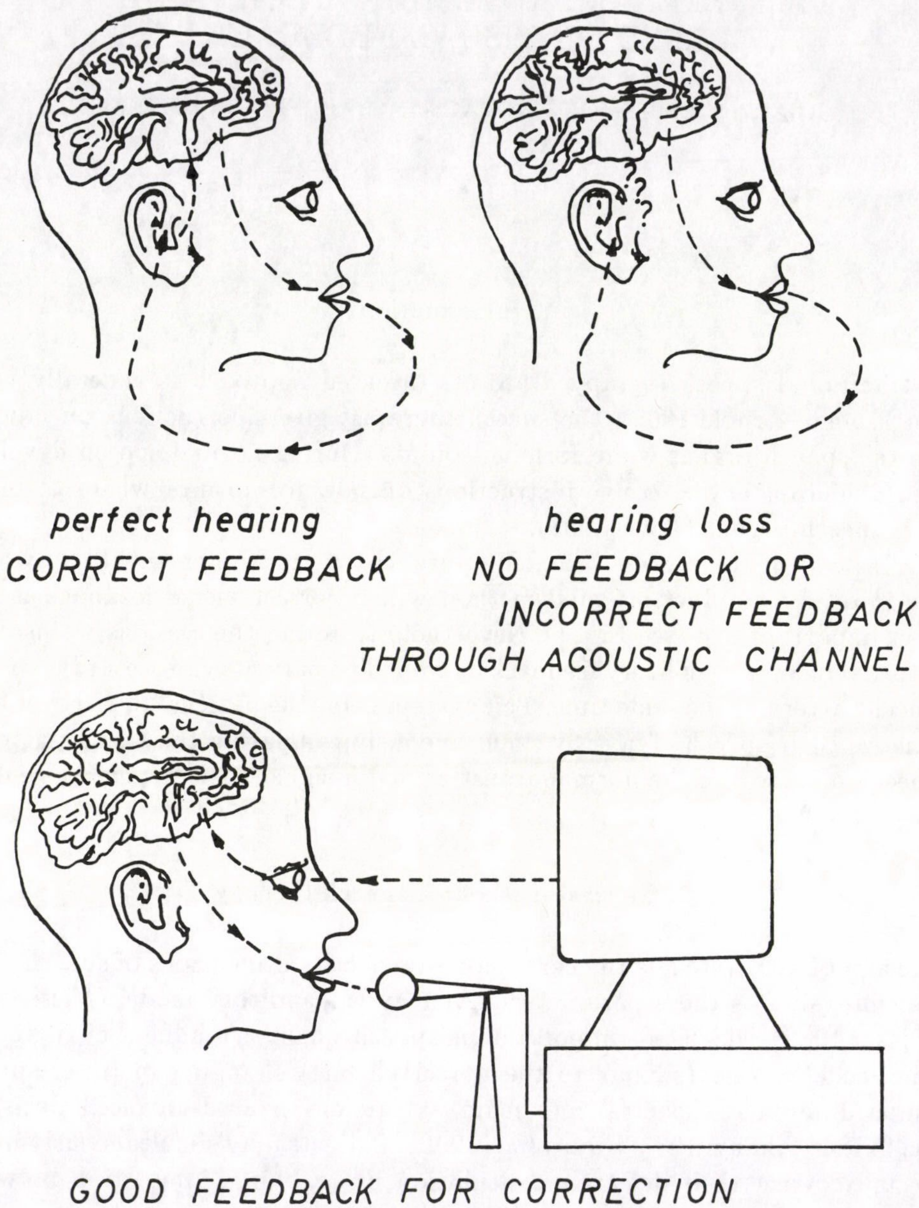


Fig. 1
General concept

One possibility is to practise the sound in isolated pronunciation, where the spectral distribution of the correct articulation is represented by a funny drawing. The measured data must fall within the allowed distribution line. The other possibility is to practise the sound in different sound combinations like syllables, words, and sentences.

In sound combinations, the patients see how the different parameters change in time, and they can compare their pronunciation with the correct one as the upper half of the computer screen shows the correct pronunciation (the etalon), while in the lower part the child's imitation can be seen.

3. The steps of the therapy

Our system is new in concept, and is based on up-to-date technology, but we follow the steps of traditional speech therapy. These are the stages of sound preparation, sound development, practice in words and automation.

3.1. Sound preparation

In the sound preparation stage we turn the child's attention to the computer screen where he observes the changes of movement in speech organs. Through this method the child is able to comprehend the fact that each sound is associated with a definitive, characteristic spectrum (speech picture). For example, it is demonstrated how the patients get different speech pictures while making a continuous air flow and moving the tongue forward or backward.

3.2. Sound development

The development of a sound starts by showing the symbolic picture used in traditional speech therapy and the written form of the phoneme together with a simplified drawing of the position of the speech organs. An example is shown in Fig. 2 for [z].

3.2.1. Sound development in isolated pronunciation

The isolated pronunciation of the sound is practised first in most cases. On the computer screen the acceptable distribution of spectral components of the sound is visualised in a playful form, as it is shown for the [ʃ] sound in Fig. 3. For instance, in the case of [s], the average spectrum of correct

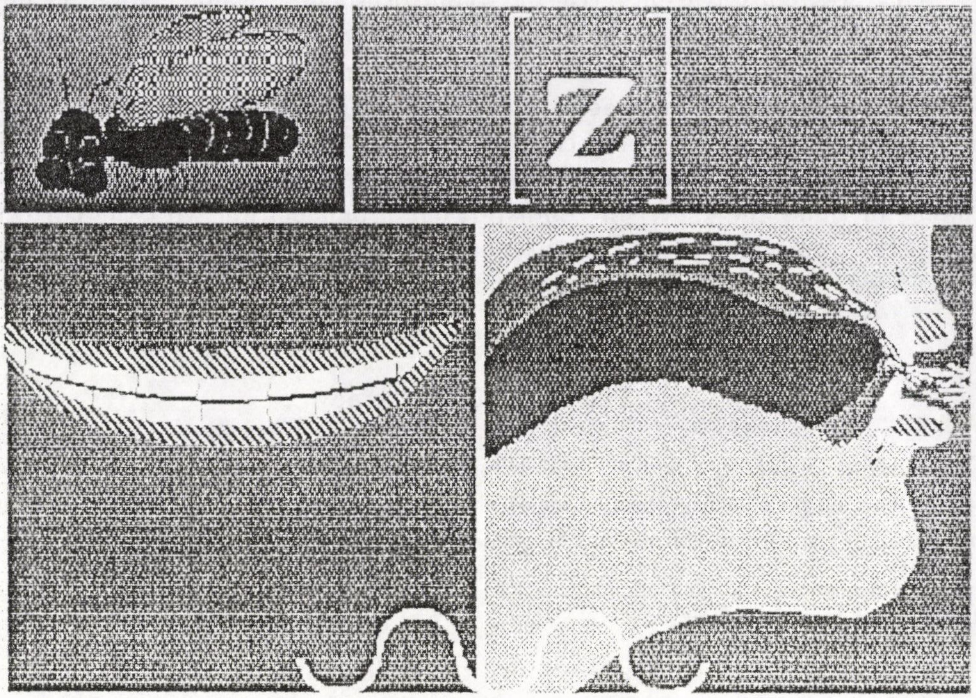


Fig. 2
A simplified drawing of the position of the speech organs for [z] together with the symbolic picture used in traditional speech therapy and the written form of the phoneme

articulation is represented by the dividing line of a yellow road, while the road itself characterizes the allowed range of variation. In response to correct pronunciation the spectra will remain within the road and follow the line of it. For [ʃ] the course of a river shows the correct pronunciation (usually the child is asked to blow waves on the river). The concept is the same for all the sounds which are practised separately. The measured data will be outside the allowed distribution line if the pronunciation is not correct.

3.2.2. Sound development in sound combinations

Sound development includes the practice of sounds in syllables. In that case, the child is confronted with the change of the spectra in time, and the pronunciation is seen as a complex speech picture. There are a few instruments in use

Fig. 3
Expected distribution of spectral components of [ʃ]

which show spectral changes in time, but the picture is too complicated for children. In our system, only the important part of the cochleogram is visible, and the speech picture is quite simple. Nevertheless, it is still too difficult for children to interpret. Therefore we try to teach the children how to interpret the cochleogram. This speech picture is shown together with various funny drawings. The attention of the child is attracted to the important element of the speech picture by these drawings and by playing with colours.

The sounds currently practised can be emphasized by red colour on the computer screen, moreover, the drawn picture symbol of the sound (covered by the cochleogram) helps the child to learn which part is important in the cochleogram, and which part is less important. An example is shown in Fig. 4 for the [ʃ] sound. In case of correct pronunciation the smoke from the train has to be covered and the train itself has to remain visible. Through continual practice, the child can easily adapt to reading the spectra (of course, only as exactly as needed), and when he observes the distorted pronunciation he will have the ability to correct it himself.

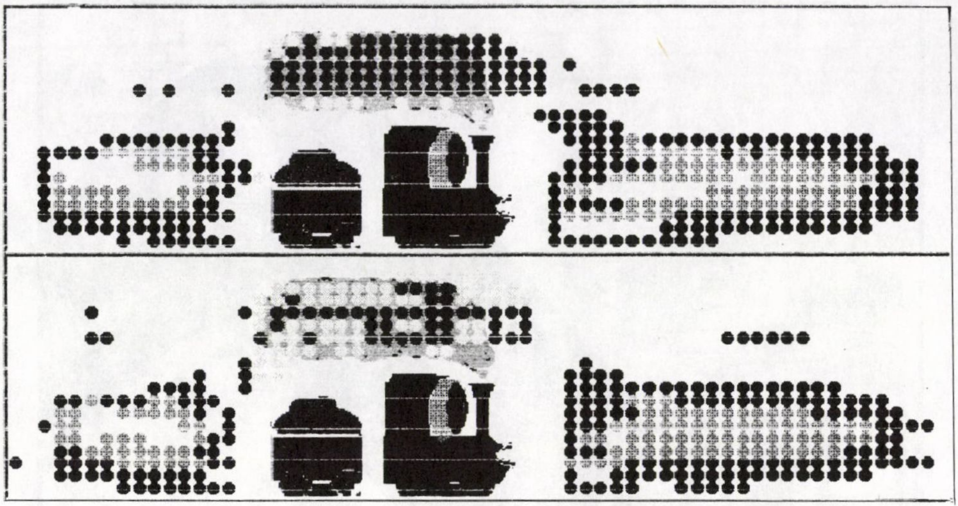


Fig. 4
The cochleogram of the [ʃ] in the sequence [ɔʃɔ]. In case of correct pronunciation the smoke from the train has to be covered and the train itself has to remain visible

Different vocabularies exist to show the cochleogram of the clear and correct pronunciation of the sound in different positions. The sounds are practised in several syllables, the target sounds being in the beginning, at the end, and in the middle.

3.3. Practice in words

The basic method here is the same as that described in the section on "sound development in sound combinations". Separate vocabularies exist to practise the sound in the beginning, at the end, and in the middle of the word. The words can follow each other in a definite didactical course or the selection can be optional. The practised sound exists in all frequently used combinations in the vocabularies.

With young children, age 4–6, the funny drawing is used together with the speech picture.

With adults or older children, another possibility exists. A mathematical distance computation is done between the cochleogram parameters of the correct pronunciation and the imitation by patients. Light green colour shows if the distance is too large between the etalon and the imitation.

3.4. Automation

Our aim in the therapy is to reach a level where the patient speaks correctly even when he does not concentrate on articulation.

In our system they can practise the sounds in sentences, too. Special sentences, first simple, then complex ones, have been collected in the vocabulary.

4. Results in the correction of sigmatisms

Our method was first applied to correct sigmatism. Sigmatism involves all kinds of defective pronunciation of sibilants. Those are the most common forms of articulation disorders. Otherwise, in high frequency hearing loss the child may hear sounds like [s] and [ʃ] so faintly that they are virtually nonexistent to him, or even if he hears these sounds, he may hear them distortedly so that they cannot serve as a model for correction (van Riper–Emerick 1984).

Certain cases are rather resistant to therapy. Children need more and continuous practice. We kept this fact in mind and the therapy program was worked out both for special teachers and for home training. It is adequate for different age groups (i.e. for children and adults), moreover we used it for cases with various etiological backgrounds, such as hearing problems, cleft palate, disorders of auditory perception, etc. (Daniloff *et al.* 1980; Krech 1954, 1969).

On the basis of acoustic-phonetic research of sigmatisms we determined the characteristic features of these disorders (Kovács-Vass 1983). We used the most characteristic types of sigmatisms (interdental, strident, lateral) which differ from normal sound production in a definite way. The interdental type is pronounced by low frequency components; that means the sharpness of the sound is lacking. In the case of strident sigmatism the sound is too sharp. Lateral sigmatism is the most offensive and unusual form. It is characterized as 'slurping' or a 'hushy' sound.

On the basis of many experiments the allowed spread of the sounds was established for isolated normal pronunciation, and the place and form of drawings, which help to read the important part of cochleogram, were found out in sound combinations. All were done so that they show the mistakes clearly, but do not reject the successful attempts.

The therapy was given to 8 speech handicapped children: 2 of them have normal hearing, and 6 of them have strong hearing loss especially in high frequencies.

Our findings were the following:

1. Children can use the system very easily and they use it with pleasure.
2. The sibilants were formed sooner than in control groups.
3. When these sounds were very resistant to the traditional therapy, the new method helped to prepare these sounds.
4. Visual feedback helped the children to see whether their pronunciation was correct or not, and how far they were from the correct one. They did not have to rely only on the teacher's opinion. In particular, this is important in cases of a speech handicap with hearing loss.

Of course, especially for small children the visual tool itself does not substitute for the work of the speech therapist. The tool is a good aid. It helps the work of the therapist and gives variety to teaching. On the other hand, after automation, or in the case of youngsters and adults, the visual tool itself gives the possibility to practise alone.

References

Daniloff, R.G.–Wilcox, K.–Stephens, M.I. 1980. An acoustic-articulatory description of children's defective [s] productions. In: Journal of Communication Disorders 13: 347–63.

Kovács-Vass, E. 1983. Experimental Phonetic Research of Oral Sigmatisms. Ph.D. dissertation. Prague.

Krech, H. 1954. Sprechkundliche Beiträge zur Therapie der Sigmatismen. In: Wissentschaftlichen Zeitschrift der Martin-Luther-Universität 5: 833–86.

Krech, H. 1969. Die Behandlung gestörter S-Laute. Volk und Gesundheit, Berlin.

Luchsinger, R.–Arnold, G.E. 1965. Voice–Speech–Language. Wadsworth, Belmont.

Povel, D.J. 1991. The visual speech apparatus: theoretical and practical aspects. In: Speech Communication 10: 59–80.

Van Riper, C.–Emerick, L. 1984. Speech Correction: An Introduction to Speech Pathology and Audiology. Prentice-Hall Inc., Englewood Cliffs, N.J.

Vicsi, K. 1986. Automatic word recognition system using some results of psychoacoustics. 6. FASE Conference, Sopron, 3.A.8.

Vicsi, K.–Mattila, M.–Berényi, P. 1990. Continuous speech segmentation using different methods. In: Acustica 71: 152–6.

Vicsi, K.–Tihanyi, O. 1987. A version of the speaker dependent isolated word recognition systems. Joint Seminar on Speech Processing, 31–4. Budapest.

Zwicker, E. 1982. Psychoakustik. Springer Verlag, Berlin.

Address of the authors: Klára Vicsi and Péter Barczikay
'György Békésy' Acoustic Research Laboratory
Budaörsi út 41–43
1118 Budapest
Hungary

Emőke Kovács-Vass
'Gusztáv Bárczi' Teachers' Training College
for Handicapped Children
Bethlen tér 2.
1071 Budapest
Hungary

Acta Linguistica Hungarica, Vol. 42 (1-2), pp. 103-115 (1994)

COGNITIVE LANGUAGE THERAPY IN GLOBAL APHASIA

ÁGNES HEGYI

1. Theoretical bases of the therapy

1.1. Aphasia is a disorder of **language ability**, defined as follows.

A human being as a superior creature due to his anatomic and neurophysiological properties possesses an ability different from other living creatures, which—with the help of the second signal system—makes it possible for him to perceive and comprehend information coming from the world in a way that he uses the verbal symbols accepted by the community according to grammar rules and supplements it with nonverbal symbols. Language ability is not only the condition of a successful language activity but it develops during this activity.

Our definition of language ability defines the aim of aphasia therapy. During the linguistic restructuring our aim is to enable the patient to understand information and produce it. So the patient uses again the conventional symbols of the language community according to the accepted grammar rules and he supplements them with nonverbal symbols in all the modulations of the language. It is suggested by the foregoing that the therapies restructuring the language abilities should be different from the therapies applied in phonation disorders and speech disorders (Lundman–Tenenholtz–Galyas 1978; Hegyi 1988).

1.2. If we want to understand the problems involved in aphasia in general—starting out of our definition—on the one hand, we have to delimit it from the other language disabilities. On the other hand, we have to define aphasia and the other cerebral disfunctions closely linked to it but distinguishable from it: perception, gnosis; execution, praxis; as well as writing, reading, speech and gesture (Fig. 1).

On the input side of the communication model the stimuli strengthening one another reach the central nervous system through the primary perception channel. The recognition of information may take place after the different stimuli are perceived. According to the models of linguistic processing of

1216-8076/94/$ 4.00

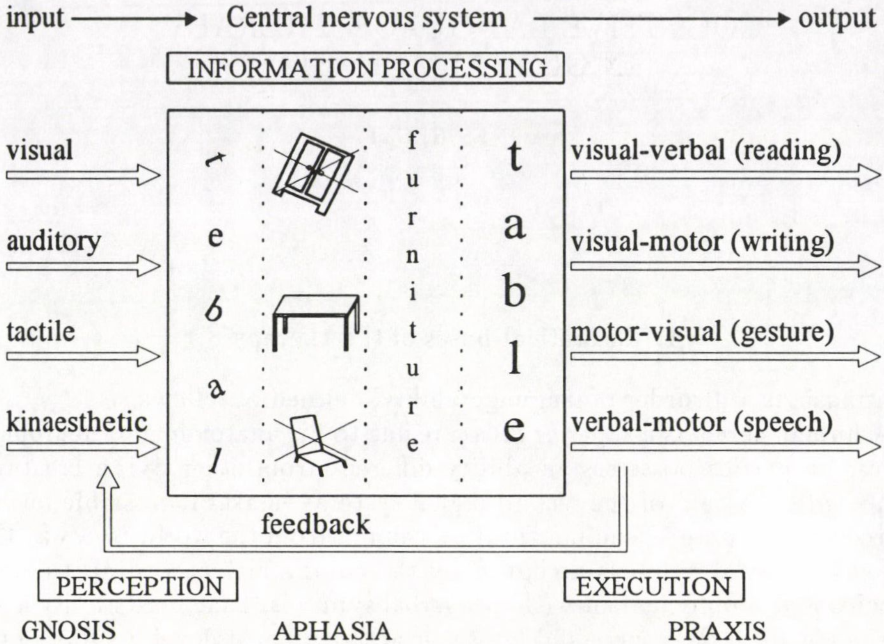


Fig. 1
Information processing at word retrieval level

information the adequate semantic, phonologic, phonetic as well as graphic representation is presumed by the process of conceptual construction. The execution is made possible by their joint effect through several channels on the output side. Feedback as a control process may prove the accurate execution, which can be realised by means of reading, writing, speaking or in deaf people by "sign language" replacing speech (Poizner–Bellugi–Klima 1990).

1.3. The cause of deficient language ability is to be found in the deficiency or disorder of information processing of the nervous system (Marshall–Newcombe 1973; Newcombe–Marshall 1981). An impairment at some point of information processing hinders the patients from retrieving the language symbols according to the rules of language and from applying them as a means of communication (Finkelnburg 1870; Froment 1921; Head 1926; Helm-Estabrooks–Fitzpatrick–Barresi 1981; Howard *et al.* 1984). The main task is to help the patients in retrieving the language symbols—at all language levels—so as they could construct the correct sequence of them in accordance with the patients' intention.

1.4. The disorder of symbol usage in reading and writing appears in the same transmission way as in the disorder of expression and comprehension (Poeck 1982; Lang–Stockert 1986). The knowledge of information processing in reading and writing helps reconstruct the language ability of aphasia patients. So our task is to reconstruct simultaneously the linguistic modulations strengthening one another—writing, reading and speech—in which gestures, the movements of hands and body have a natural role (cf. Shewan–Bandur 1986).

2. Patient selection criteria

In selecting the patients for therapy we relied on the clinical (neurological and internal medical) diagnoses and the results of functional tests. Linguistic, psychic and psychosocial factors were taken into consideration beside the neurological factors. The neurological diagnoses were accurate due to the following: the local impairment of the brain or that of the dominant hemisphere from the point of view of speech was proved by the results of the paresis examination, reflex test, electrophysiological examinations (EEG, EMG) and cerebral arteriographical examination (AG). All the treated patients were right handed and the left hemisphere was impaired.

EEG (Electro-encephalography): The electrodes fixed on the skin of the head (according to the area of the brain) send electric waves about the qualitative and quantitative function of the brain, and their graphical signs give appreciable information about the disorders of the brain.

EMG (Electro-myography): An electric device for describing voluntary muscle activity. Normal and pathological muscle activity can be distinguished by its help.

AG (Angiography): After the cerebral vascular system is stained with contrast material, it can be seen on an X-ray plate. The photograph shows the blockage and deformations of the vessels.

The brain impairments were of vascular origin which caused a large blood supply deficit in the fronto-temporal and parietal lobes with the structural degeneration and functional disorder of these areas. The examination of the visual field showed the intact character of the posterior occipital lob and optic nerves. Because of this latter criterion the application of our therapy relying on the intact visuality was possible (Helm–Barresi 1980). Stability in the neurological and general physiological condition are expected of the patient, that is why we took into account the case history, prodromal symptoms of stroke, hyper tension, vascular spasm, numbness of the limbs, and other basic illnesses associated with the given vascular condition; diabetes, heart diseases, peripherial vascular damage.

From the point of view of the language it was important for the patient to show readiness for verbal contact. The simplest nonverbal signs (glance, touch of the hand) are unambiguous indicators. The possibility of primary hearing impairment is excluded when the patient reacted to the noises of the surroundings and to the sounds of speech.

The ability to produce basic voices (voice reaction induced by a spatula) gave information about the fact that the patient was not anarthric, that is, his sound forming organs were intact, the pathways for articulation movements could be innervated. The ability to reproduce the rhythm even in a dermo-tactile way was important from two points of view. From a linguistic point of view this ability showed that the patient would be able to syllabify on the basis of internal rhythm and to maintain linguistic segmentation.

From the point of view of psychology, the ability to reproduce rhythm showed the intact character of the short term acoustic memory. The ability to retain long term memories, the intact intellect, the demand motives to cooperate with us as well as exemption from more severe psychiatric disease were important for us in the same way.

The language treatment lasting for four years always required the support of the family and the help of the surroundings. From a psychosocial point of view it was important for the patient to be near our therapeutic centre.

3. The exposition of the therapy

Figure 2 helps to survey the order of the therapeutic exercises. It shows the main groups of the exercises, the time used for them, the frequency of therapeutic sessions and their individual or group character.

3.1. In the **activating phase** the psychic functions are activated in order to help linguistic recognition with acting together, with actions accompanied by strong gestures and verbal directives associated to them. In the patient's repeated object activity the visual, auditive, tactile, and kinaesthetic senses are arranged into perceptions. Perceptions repeated in an analogous way help recognize the object and operation of the action. Merdian (1984) proved that the functions of linguistic recognition and naming are improved by developing the preverbal learning ability.

Treatment frequency	Time period from-to		Stages	
everyday	2–4 weeks		ACTIVATION PHASE	
everyday	4–12 weeks	individual	SYNDROME-SPECIFIC EXERCISES	Sound retrieval level
3/week	8–52 weeks			Word-structure level
2/week	10–24 months			Sentence-structure level
2/week	1.5–3 years			Text structure level
1/week	3–4 years	group	CONSOLIDATION PHASE	

Fig. 2
Structure of therapy

3.2. In the **syndrome specific exercises** we develop the abilities that help the patient regain his language ability. The exercises are connected to the various language levels. In the "starting phase" of language therapy—at the level of word formation and sound retrieval—our aim is to establish the writing and reading ability, that is, to trigger the function of cognitive information processing by means of linguistic analysis, synthesis and, parallel synthesis–analysis (cf. Pavlov).

In the "advanced phase" of therapy—at sentence and text forming level—speech is improved because the information processing in reading and writing offers a permanent external control for phonology, grammar and semantics, in contradiction to speech, where control of this type should be of an internal character (Lytton-Brust 1989).

If in the activating phase the patient is able to select first colours, simple geometrical shapes, that is, "prototypes" (Ádám 1987) then arranging objects into identical groups on the basis of several properties, his memory is able to register and retain several stimuli simultaneously. This ability must be utilized so that the patient could retrieve language symbols connected to sound formation, to create a simultaneous link between the articulated sound and the heard sound and the printed letter. In our opinion the retrieval of vowels as musical elements is promoted by the intact operation of the subdominant hemisphere (Pléh 1981; Lebrun 1983; Geschwind–Galaburda 1985). The sequence of retrieval of Hungarian vowels is shown in three groups in Fig. 3.

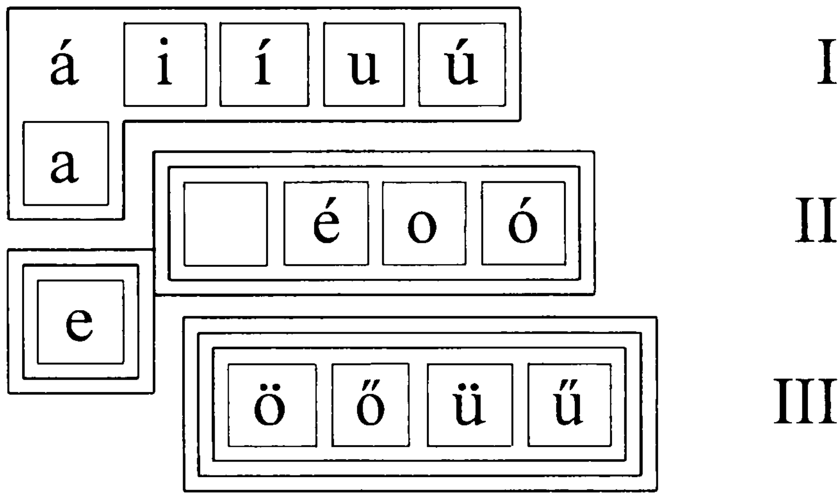


Fig. 3
Sequence of retrieval of Hungarian vowels

When distinguishing the vowels and taking them out of the sequence of sounds, auditive, visual, kinaesthetic and tactile stimuli are applied. The therapist and the patient articulate the sound while showing the colour backgrounded printed form of the letter, then accompany their joint sound pronunciation with phonomimicry and the sound retrieval is even supported by the "code-word" belonging to a particular sound (de Partz 1986). The patient identifies the sound by the corresponding letter. Then the patient is asked to write and utter it individually.

The retrieval of consonants and their differentiation from other sounds in the mother tongue starts after the correct recognition of the first four vowels. Similarly, as in the case of vowels, here we also apply the method of indirect stimulation (Weigl 1979). Even when the first four vowels and only one consonant are retrieved by the patient, from his poor stock of sounds and letters, he will be able to form a sequence of sounds and letters for a real word.

Word formation is understood to be a visual building of a word percieved previously in an acoustic way.

The retrieval of sound and letter symbols of a whole word uttered acoustically and the forming of their corresponding sequence of sounds is taking place first in groups formed according to their formal properties. The word forming is supported by the therapist's strategic orders which refer to the place of change in sound or letter and what kind of change can be expected (Fig. 4). For instance, "Listen to the beginning of the word, it will be one letter longer!". The retrieval of the symbol value of the word is supported by the recognized minimal phonological difference and the appearance of the word in situational context, that is, in a sentence (Howard *et al.* 1985).

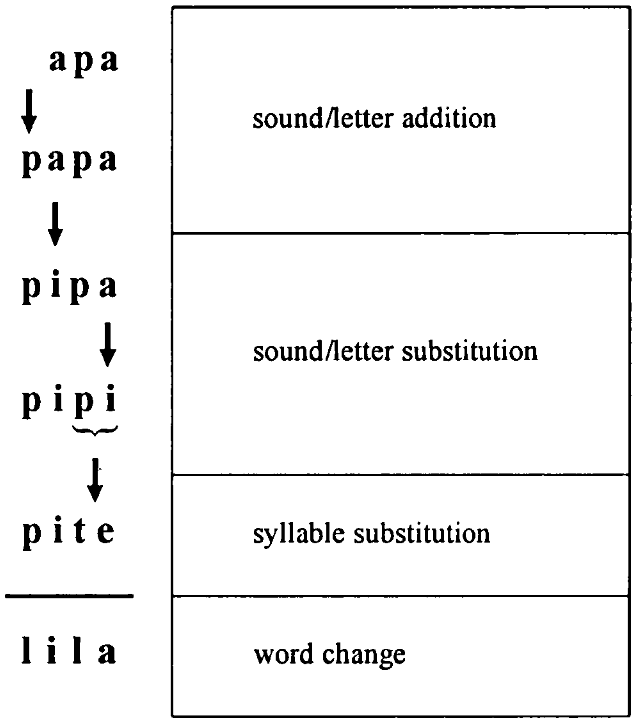


Fig. 4
Formal factors on word formation level
Groups of formal properties

For every session the therapist chooses the group of words by their form properties according to the patient's sound retrieval ability. That is why a collection of pictures concerning the words cannot serve to be a long lasting means for supporting the semantic content of the word. Instead, the words (e.g. their etymological derivatives) are presented in sentences or in part of a sentence and the therapist picks the word to be put together out of the context and later the patient does the same.

The patient is more and more unaided in word retrieval within the groups defined by the meaning of the word. The content factors of words at the word formation level are constructed by the groups of meaning according to the following grades: by antonyms (*small*, *large*), synonyms (*dog*, *hound*), semantic field (*furniture*, *table*, *chair*), partial and total connections (*table*, *drawer*), functional connections (*plane*, *wood*, *saw*, *glue*), and situational context (*booking office*, *train*, *ticket*, *conductor*, *refreshment car*, *newspaper*, *railway station*).

In functional connection and situational context the objects appear on the basis of the patient's knowledge linked to the action. In our opinion it is correct if we continue the exercises at the sentence forming level. Here sentence formation as expressed in speech is realized through writing and reading. For this work the Hungarian author Ildikó Meixner's textbook for dyslexic children is used.

During the same therapeutic session we get to reading the sentence through reading letters, syllables and words. We make the patient read the words by selective limitation that is in dependence of the reading channel he uses. We make him repeat the red word and make him explain it. In the explanation of words both expressions of "verbal" and "nonverbal" means can be accepted (Davis–Wilcox 1985; Glindemann–Springer 1989).

The forming of simple, complex and compound sentences are supported by analogue sentences, through which empirical generalization is possible on the basis of surface properties (word order, suffixes, etc.) without giving names to the grammatical categories.

When practising the order of tenses the patient arranges the sentences into sequence of events, that is, into a text. We use the picture stories made by the Austrian Friederica Meixner (Fig. 5).

Written questions concerning the individual pictures of the story are asked. In the question the patient perceives the elements connecting the text (directives, anaphoras, conjunctions etc.). In his oral answer, then in a written form, he compulsorily uses them. With their use the patient will be able to adjust the next sentence to the previous one.

Text retrieval after reading helps in developing reading and text forming at a higher level. In the poetic translation of La Fontaine's fables there are not many redundant elements, but there are far more expressions with new information, which do not support the patient's guessing on the stereotype

Fig. 5
Picture story at text forming level

expressions (Fónagy 1962). It is possible to replace the cited dialogue by the grammatically more complicated oblique oration when the fable is put down and retold.

3.3. In the **phase of consolidation** the aim is to practise the ability of general language communication in a situation created by the patient. In group therapy the information exchange is mutual after reading a literary work and the patients are equal communication partners in the discussion. The use of language is motivated by conversation (Cvetkova 1980; Glocman 1981).

4. Results and conclusions

Figure 6 gives information about results of treatment of patients suffering from global aphasia who participated in the program of cognitive language therapy in the period of 1982–1990.

Nineteen aphasia patients diagnosed by neurological and functional examination to be global aphasics were treated. There were seven female and

Number of patients Female Male	19 7 12
Age range (years) Mean age (ys) ±SD	33–70 52.9±12.5
Education	mainly high level
Neurological disease	acute cerebral vascular accident (stroke), left hemisphere
Type of aphasia Associated neurological symptoms	global right hemiparesis right hemihypaesthesia right hemiparaesthesia
Average treatment period	at least 4 years i.e. 420 working hours/person
Factors in unsuccessful treatments (n=4)	low education level (or earlier illiteracy) low intellectual capacity inability of keeping rhythms: – disturbance of immediate memory – inability of linguistical segment (syllable)

Fig. 6
Summarizing table about the treated patients and the results

twelve male patients registered. Their average life time was 52.9 years. Beside the younger and older five patients the majority was betwen 50 and 60. The average treatment time was four years which means that a successful treating period needed 420 hours.

15 of our patients were treated with good result. They are able to carry out consciously information exchange in speech, reading and writing in any communicational situation.

The treatment of four patients failed. In these patients the linguistic, psychic and psychosocial parameters proved to be unfavourable. Their failure can be explained by the following factors:

(a) Low educational level and the low level of intellectual receptivity prevents the accomplishment of work in which we wish to reconstruct the language ability of the patient or retrieve it as a function of the cognitive abilites in the patient.

In vain did we teach an illiterate woman, 70, to recognize the sounds/letters, to write them. Similarly global word retrieval was not supported by word-picture reinforcement. Finally the limited number of word retrieval with the help of object pictures did not give possibility to restructure the language ability in a system. Another 70-year-old female patient had completed only four classes of elementary school, was a housewife and lived with her family in a farm. The patient did not possess the intellectual receptivity and demand to succeed in reading and writing language therapy.

(b) The rhythmical knocking disability exerts an unfavourable influence on the success of our therapy from two points of view. From a psychological point of view we may presume the impairment of short term memory, whereas from a linguistic point of view we may expect the loss of syllabification ability on the basis of internal rhythm.

Two male patients, 58 and 62 respectively, were not able to knock a rhythm and we were not able to teach them to do it. Both patients were referred to us and we began the therapy with them a month and a half and two months after the impairment. By that time they had developed a jabbering, unintelligible speech full of neologisms. Their ability to maintain a verbal contact was unilateral, with the intention to say something but failing to comprehend their partner. Due to the lack of forming a lasting contact our therapeutic endeavours proved to be useless in the activating phase.—The similarity of the two cases leads to a conclusion that the blockage of the medial cerebral artery is more extended in both patients in comparison with the treatable patients and these caused an acute verbal and constructive apraxia. It would have been worth making a computer tomography (CT) about the extent of the impairment to make it certain but the device has been accessible in the Clinic of Radiography, Szeged, only since the early 1990s.

(c) In the two latter cases the fact must be kept in mind that the patients received only medical treatment for a month and half or two months and they faced the lack of understanding among the family members. Patients improving during the therapy took an active part in an aimed therapy since the second but no latter than the fourth week of their disease. In our opinion the patient's recovery is influenced by the time between the beginning of the illness and that of the therapy (Darley 1982; Reinwang 1984; Lang-Stockert 1986).

A conclusion drawn from our results and failures is that the therapy described in this paper, with special emphasis on linguistic considerations, makes complete recovery of language ability of the global aphasia patient possible. This linguistic cognitive therapy, however, shows further perspectives beyond the application described. We may have an exact result about the work of language information processes in reading, writing and speech in any type of aphasia after diagnosing the patient by a standardised linguistic test; e.g. Boston, Western or Aachen Diagnostic Aphasia Test. We suppose that the patient's disorder in his language information processes can be improved by dropping the deficiency symptoms if the reeducation is realized by activating the cognitive processes necessary for the functioning of the language system.

References

Ádám, Gy. 1987. A megismerés csapdái [Pitfalls of cognition], 45–71. Budapest, Magvető Kiadó.

Cvetkova, L.S. 1980. Some ways of optimisation of aphasics' rehabilitation. In: International Journal of Rehabilitation Research 3: 183–90.

Darley, F.L. 1982. Aphasia. W.B. Saunders, Philadelphia–London–Toronto–Mexico City, Rio de Janeiro–Sydney–Tokio.

Davis, G.A.–Wilcox, M.J. 1985. Adult Aphasia Rehabilitation: Applied Pragmatics. NFR-Nelson, Windsor.

Finkelnburg, F.C. 1870. Asymbolie und Aphasie. In: Berliner Klinische Wochenschrift 7: 449–50, 460–2.

Fónagy, I. 1962. Beszéd és valószínűség [Speech and probability]. In: Magyar Nyelvőr 86: 309–20.

Froment, J. 1921. La reeducation des aphasiques moteurs: principes, procédés et résultats. Paris, Medical 41: 267–73.

Geschwind, N.–Galaburda, M.A. 1985. Cerebral lateralization. In: Archives of Neurology 42: 428–59.

Glindemann, R.–Springer, L. 1989. PACE-Therapie und sprachsystematische Übungen. Ein integrativer Vorschlag zur Aphasietherapie. In: Sprache-Stimme-Gehör 13: 188–92.

Glocman, J.M. 1981. On increasing motivation to communication in aphasics' rehabilitation. In: International Journal of Rehabilitation Research 4: 78–81.

Head, H. 1926. Aphasia and kindred disorders of speech. Cambridge University Press, Cambridge.

Hegyi, Á. 1988. Neurolingvisztika a logopédiában [Neurolinguistics in speech therapy]. In: Művelődési Minisztérium Gyógypedagógiai Továbbképző Könyvtára 28: 152–9. Budapest.

Helm, N.A.–Barresi, B. 1980. Voluntary control of involuntary utterances: A treatment approach for severe aphasia. In: Brookshire, R.H. (ed.): Clinical Aphasiology Conference Proceedings 1980. BRK, Minneapolis.

Helm-Estabrooks, N.A.–Fitzpatrick, P.M.–Barresi, B. 1981. Response of an agrammatic patient to a syntax stimulation program for aphasia. In: Journal of Speech and Hearing Disorders 46: 422–7.

Howard, D.–Patterson, K.E.–Franklin, S.–Morton, J.–Orchard-Lisle, V.M. 1984. Variability and consistency in picture naming by aphasic patients. In: Rose, F.C. (ed.): Recent Advances in Neurology 42. Progress in Aphasiology. Raven, New York.

Howard, D.–Patterson, K.E.–Franklin, S.–Orchard-Lisle, V.M.–Morton, J. 1985. The treatment of word retrieval deficits in aphasia: a comparison of two therapy methods. In: Brain 108: 817–29.

Lang, Ch.–Stockert, Th.R. 1986. Zum gegenwartigen Stand der Aphasietherapie. In: Fortschritte in der Neurologie und Psychiatrie, 54: 119–37.

Lebrun, Y. 1983. Cerebral dominance for language. A neorolinguistic approach. In: Folia Phoniatrica 35: 13–39.

Lundman, M.–Tenenholtz, E.–Galyas, K. 1978. Technical aids for the speech impaired. Proposal for research and development activities 9: 43.

Lytton, W.W.–Brust, J.C.M. 1989. Direct dyslexia. Preserved oral reading of real words in Wernicke's aphasia. In: Brain 112: 583–94.

Marshall, J.C.–Newcombe, F. 1973. Patterns of paralexia. In: Journal of Psycholinguistic Research 2: 175–99.

Meixner, I. 1990. Én is tudok olvasni [I can read, too]. Tankönyvkiadó, Budapest.

Merdian, G. 1984. Semantische Störungen bei Aphasikern. Europaische Hochschulschriften. P. Lang, Frankfurt am Main–New York–Nancy.

Newcombe, F.–Marshall, J.C. 1981. On psycholinguistic classifications of the acquired dyslexias. In: Bulletin of the Orton Society 31: 29–46.

de Partz, M.P. 1986. Reducation of a deep dyslexic patient: rationale of the method and results. In: Cognitive Neuropsychology 3: 149–77.

Pléh, Cs. 1981. Különböző szórendű mondatok értelmezése és a dichotikus hallási aszimmetriák 3–6 éves gyermekeknél [Interpreting sentences with differing word orders and dichotic hearing asymmetries in children aged 3–6]. In: Pszichológia 3: 365–93.

Poeck, K. 1982. Leitungsstörungen. In: Poeck K. (ed.): Klinische Neuropsychologie, 159–67. Thieme, Stuttgart.

Poizner, H.–Bellugi, U.–Klima, E.S. 1990. Biological foundations of language. Clues from Sign Language. In: Annual Review of Neuroscience 13: 283–307.

Reinwang, I. 1984. The natural history of aphasia. In: Rose, F.C. (ed.): Progress in Aphasiology. Advances in Neurology 42: 13–24. Raven, New York.

Shewan, C.M.–Bandur, D.L. 1986. Treatment of Aphasia: A Language-oriented Approach. Tailor and Frances, London.

Weigl, I. 1979. Neuropsychologische und psycholinguistische Grundlagen eines Programms zur Rehabilitierung aphasischer Störungen. In: Peuser, G. (ed.): Studien zur Sprachtherapie, 491–514. Fink, München.

Address of the author: Ágnes Hegyi
Department of Neurology and Psychiatry
Szent-Györgyi University Medical School
Szeged
Hungary

Acta Linguistica Hungarica, Vol. 42 (1–2), pp. 117–145 (1994)

BOOK REVIEWS

Tamás Szende: Phonological Representation and Lenition Processes. Magyar Fonetikai Füzetek/Hungarian Papers in Phonetics 24. Linguistics Institute of the Hungarian Academy of Sciences, Budapest 1992. 269 pp.

In his new book, Tamás Szende discusses phonological processes that turn underlying phonological representations into natural surface phonetic sequences as used in everyday colloquial speech. The spoken realizations of words normally deviate from their underlying forms to some extent. Naive (i.e., normal) native speakers are usually unaware of such deviations even though they make constant use of the processes concerned. For instance, the common realization *-ba* of the suffix *-ban* 'in' is regarded by many people as unacceptable except in familiar or casual styles of speech. The insistence on pronouncing the *n* in this suffix or, more generally, the expectation that phonological representations be respected in what is referred to as 'careful speech' may have rather embarrassing consequences. Thus, Zoltán Latinovits's acclaimed recitals of poems by Ady or Arany could be dismissed either as not belonging to the realm of 'careful speech' or else as unsuccessful attempts at speaking correctly. Both claims are absurd.

In order for Szende's enterprise to be feasible, two preliminary requirements had to be fulfilled: phonological representation as an abstract–general category had to be defined and a reliable and authentic data base (cf. Ilson 1991) had to be compiled.

The data analysed in the book were taken from a spontaneous conversation of four college students, aged between 21 and 23, and a 43-year-old engineer. The recording session took place in 1986 in the phonetics laboratory of the Linguistics Institute of the Hungarian Academy of Sciences. The total speech output produced by the three female and two male subjects ran into approximately 45,000 syllables. The material was apportioned into 2,055 samples, each containing 1 to 20 lenition cases (or, distortions of phonological representations). A carefully selected subset of these samples was then submitted to instrumental analysis.

In the first chapter (Theoretical approaches to word-level phonological representation in post-SPE frameworks, 9–86) the author presents a critical overview of theories put forward between 1968 (the publication of Chomsky and Halle's "The Sound Pattern of English") and 1987: he discusses a total of twelve different frameworks from natural generative phonology to dynamic phonology. The final conclusion emerging from this detailed but not over-meticulous critical survey is that none of the phonological theories discussed is able to provide a full-fledged and practicable framework for the description of word-level phonological representations that would suitably set the stage for an account of lenition processes (86).

But what is **lenition**? In terms of natural phonology, each sequence has a rule-governed realization strictly corresponding to its phonological representation that will, under certain conditions, undergo various **distortion processes** during speech production. Instances of **fortition** are cases in which the realized sequence of segments is more articulated than the phonological representation; e.g. assimilations are suppressed as in *út-ja* 'its way' for *úttya*. On the other hand, **lenition** takes place where the output is less articulated than the phonological form is; e.g. via the elision of some segment(s). As the author points out elsewhere: "In everyday speech these latter types of distortion constitute an overwhelming majority" (Szende 1990, 183). The distorted word forms of colloquial speech are derived from the underlying full phonological representation (where 'full' means 'containing all phonological components in a regularly ordered manner') by way of a specific mechanism. The author refers to that mechanism as "the principle of global programming" (cf. 168–181). In particular, it removes some components from the full representation and, within certain limits, disrupts its ordered character. The result is a narrowed information field: a partial restructuring of the word form in the distorted realization. (This explains why casual speech as a genre is more difficult to comprehend than formal speech is.)

The second chapter (Systematic phonological representation, 87–183) clarifies issues that must be settled before processes leading from phonological representation to surface realization can be adequately accounted for. These issues include a reappraisal of the concept of **phoneme** that had been eclipsed in classical generative phonology (Chomsky–Halle 1968), as well as the effects of Szende's notion of phoneme on the description of Hungarian vowels and consonants. A number of currently debated points about Hungarian phonology are raised again, including the status of *j* (whether it is a consonant or the onglide/offglide of a diphthong), that of long consonants (whether they are independent phonemes or sequences of identical short consonants), the place of mid *ë* (as opposed to low *e*) within the vowel system of Hungarian, and so forth.

The third chapter (Lenition processes, 184–218) first defines the concept of 'lenition' and then goes on to describe the distortion phenomena observed in Hungarian casual speech. Lenition and fortition are interpreted relative to one another. The same realization may be characterized as lenited, fortitive, or neutral, depending on the overall lenition program of its environment (190). The reasons underlying lenition in casual speech include communicative genre, (low) semantic load, as well as (increased) speech rate (higher number of phonemic components conveyed per unit of time).

The types of lenition are as follows:

1. **Reduction**. One or several articulatory gestures are omitted; one or several secondary distinctive features are incompletely implemented; syllable structure remains intact. E.g. in the *z* of *ezeket* 'these-acc.' the front of the tongue gets into a loose, partial contact with a smaller surface of the dentialveolar region of the palate; acoustically, the noise components will be shorter in duration and of lower intensity; and phonologically, the implementation of the feature [+continuant] is deficient. Its high frequency of occurrence is probably due to the fact that it involves the smallest increase of distance between phonological representation and realization.

2. **Deletion**. The primary and secondary articulatory components are eliminated from the phoneme realization, leaving phonetic traces behind; e.g. the second *n* of *nagyon jó* 'very good' may be deleted but nasalization shows up on the preceding *o*.

3. **Loss and truncation**. In the first of these, a phoneme is unrepresented in the word form. The second involves iterated elision that leaves behind at least one element of the total sequence, e.g. *szóval* 'so' $\longrightarrow$ [[so]] (where double square brackets enclose directly observable, potentially distorted phonetic representations).

4. **Reduction over the sequence**. The lenition of a longer articulatory sequence by cumulative appearance of reduction on a number of its constituents that are similar in some respect, e.g. a set of closure eliminations in *szóval egy kicsit* 'that is, a little' involving, in addition to *k* and *t*, the stop phase of the voiceless affricate as well.

5. **Sequence size truncation**. A semantically/grammatically/pragmatically identifiable portion of the sequence endures truncation of several of its constituents where realizations that are not originally part of the phonemic representation may also appear. E.g.

szóval 'so' ⟶ [[s˙]] where four phonemes are unrealized but the additional length feature appears as partial recompensation.

6. **Fusion**. Two adjacent segments form a unified articulatory/acoustic pattern that differs from the usual realization of both input segments, e.g. *szóval azt* 'so that-acc.' ⟶ [[so˙vɑ(˙)st]]. Syllable structure is affected; the number of syllables is modified depending on the syllabicity of the segments involved.

The final chapter (Systemic consequences, 219–251) discusses, among other topics, the relationship between lenition and speech rate. The latter notion is ambiguous: the rate of speech may be relatively fast even if articulation is slow, provided that several consecutive phonological units are articulated simultaneously (220). Consequently, the number of segments realized per unit of time (minute or second) is insufficient as an indication of tempo.

Lenition may occur under decreased tempo as well. E.g. in *vagy ezt kell, vagy azt kell* 'you must do either this or that' there may be *t*-deletion and *z*-devoicing "showing that some lenition processes are part and parcel of Standard Hungarian pronunciation programs in that they are quite regular even in slow or deliberate speech" (224).

Lenition affects semantically 'empty' expressions more intensively than their full (original) counterparts that are usually undistorted, cf. *szóval* 'so; that is; in a word' (adverb) vs. *szóval* 'with a word' (case-marked noun).

Consider some possible realizations of *miért* 'why' (250):

phonemic representation:		/mie:rt/
lento pronunciations:		[mie:rt], . . . , [mi^{j}e:rt]
allegro pronunciations:		
(i)	by *i*-reduction:	[m^{i}e˙rt]
(ii)	by *i*-deletion:	[me˙rt]
(iii)	by *t*-loss:	[me˙r̥]
(iv)	by fortition:	[me˙r^{ə}]

The English translation is of good quality. Péter Siptár not only translated the text into readable English but also contributed to the contents of the book during sessions of heated debate with the author (8). Nothing like the cumbrous and often ridiculous 'English' of "The Hungarian Language" (Akadémiai Kiadó, 1972), the present text is written in **English**. A small number of minor flaws do occur, though; typos and vague references could have been avoided if the volume had had a technical editor distinct from the translator.

The reader also misses an index and an accompanying tape with a fair selection of the recorded material. However, even without these, Tamás Szende's book is an important contribution to the synchronic description of Hungarian phonology.

Miklós Kontra

References

Chomsky, N.–Halle, M. 1968. The Sound Pattern of English. Harper and Row, New York.

Ilson, R. 1991. Assembling, Analysing and Using a Corpus of Authentic Language. Linguistica, Series C, 4. Linguistics Institute of the Hungarian Academy of Sciences, Budapest.

Szende, T. 1990. „Lazítási folyamatok" a köznyelvben [Lenition processes in Standard Hungarian]. In: Balogh, L.–Kontra, M. (eds): Élőnyelvi tanulmányok [Studies on the living language]. Linguistica, Series A, 3, 182–8. Linguistics Institute of the Hungarian Academy of Sciences, Budapest.

Mária Gósy: Speech Perception. Forum Phoneticum, Frankfurt am Main 1992. IX + 325 pp.

It is a rare occasion even in the wide interdisciplinary spectrum of speech perception research to have a monograph that is mainly based on original experimental data targeted on a single language. The work of Gósy is one of these rare and welcome exceptions. The main interest of the book is a summary of about two dozen experiments done mainly on sound perception, acoustic word recognition, and the perception of suprasegmentals in Hungarian. The material is well organized. It provides the reader not only with an image of high standard speech perception research but with the actual facts of Hungarian speech perception as well.

A large part of the experiments deal with the identification of isolated sounds. The author presents a large number of studies that used filtering and synthesized sounds to reveal the acoustic bases of speech identification in Hungarian. The systematicity of the studies mainly concerns the careful selection and analysis of the linguistic materials. All of this is quite reliable from the point of view of instrumental phonetics. The reader has, however, some feelings of underdocumentation with regard to the experiments themselves. We have hardly more as a task description than the cursory remark that e.g. people had to identify the speech sounds. What is missing, is a clearer image of the relationships between task dependence of speech perception itself and of the studies themselves.

A long and interesting chapter deals with word recognition in Hungarian. Some of the findings are of a general nature (e.g. even Hungarians use context in word recognition). What is specifically interesting are the studies done by the author on sentence perception with monosyllabic and polysyllabic words under noise. The author concludes that somehow the "endings of words are correctly identified ... when the content words showed misperceptions" (p. 156). This is straightforward enough and in my view relates to some of the most interesting aspects of lexical processing in Hungarian. Namely the issue of analyticity in morphological processing. The approach Gósy uses here has two basic disadvantages. One is that she fails to carefully support the above strong claim. Second, she forgets to relate in depth these findings with the theoretical models of contextual dependency in speech processing she surveys. In order to account for a superiority of grammatical morphemes in resisting noise one would need to combine the overall model of speech perception with a model of access plus morphological parsing.

The chapter on suprasegmentals—while presenting data on stress and melody as well—mainly presents data on a very original series of experiments: how do people who speak with different speed perceive other speakers' speech tempo? Though at the moment the multidimensional data are by far not clear, the reviewer has the feeling that this line of research is the most promising among the ones initiated by Gósy.

Two chapters deal with some aspects of speech perception development and with the application of the data on early screening for reading readiness, functional hearing problems and so on. This is a rather important part of Gósy's work but an aspect that she presented elsewhere already.

In sum, this is a monograph to be recommended for anyone who deals with Hungarian speech and for anyone who is interested in possible connections between speech perception models and grammatical processing. The author is a reliable and keen experimenter. She is also very well read into the up-to-date Western literature in any area she touches upon. However, the book has its disadvantages. These are not factors that question the source value of the book but they indicate a kind of asymmetry in the approach of the author between her experimental and theoretical interests. The experiments are somehow designed in a phonetically oriented way. While the author is interested in most cases in interactions between levels she usually treats grammar in an unanalysed manner. This leads to an omission of some very basic interpretive contexts. Just to name a few. The issue of modularity in processing is not mentioned even, the same way as connectionism. And more specifically, her very interesting studies on word recognition do not get integrated with the problem of morphological processing.

These shortcomings notwithstanding, this is a very useful and resourceful book.

Csaba Pléh

Magnus Olsson: Hungarian Phonology and Morphology. Travaux de l'Institut de Linguistique de Lund 26. Lund University Press, Lund 1992. 216 pp.

The progress of ongoing research on Hungarian phonology and morphology is signposted by a series of major monographs from Hall (1944), via Vago (1980) and Kornai (1986), to Abondolo (1988) and beyond. The most recent signpost in this series is Magnus Olsson's doctoral dissertation, now published as volume 26 of the Travaux de l'Institut de Linguistique de Lund, founded by Bertil Malmberg and currently edited by Gösta Bruce and Bengt Sigurd. Olsson's monograph is based rather closely on Vago (1980) [reviewed in the present journal by Siptár (1984)]. As the author himself points out, "In a way the dissertation is a criticism in the form of simplification of Vago's rules. His dialect has been accepted as the object of study ... [although] a few questionable data are explicitly mentioned. A book that is totally different from Vago's work has however emerged" (10). Some of the major innovations are as follows. (i) The basic formalism differs from earlier treatments. "As I realized the need for a different approach than that employing the SPE notation [as Vago does] but felt discomfort at the sight of the current autosegmental notation, it seemed necessary to devise a new phonological formalism" (17). The author does not specify what kind of discomfort he felt; indeed, his notation incorporates much of the insight that underlies current autosegmental formalisms but lacks the formal rigour and often the elegance of the latter; also, he makes free use of Greek letter variables, parentheses, cover symbols, subscripts, angled brackets, and the like. The only formal device he explicitly rejects is the use of curly brackets, replacing them by what he terms subset specification, see below. Thus, a rather eclectic and virtually unconstrained notational system emerges in which formal simplicity, despite repeated claims to the contrary, appears to be of very little concern. The "simplification of Vago's rules" that was referred to in the first quote usually takes the form of replacing sets of Vago's independently motivated rules by complex, sometimes bewilderingly complex, schemata, although admittedly fewer in number. (For instance, Vago's several rules for—intermorphemic—epenthesis are proposed to be replaced by a single rule schema (part of which is optional) that involves, in addition to multiple Greek letter variables and labelled angled brackets, also syntactic features and even inversion around the morpheme boundary in certain environments.) As a result, even the cases where Olsson's analysis is conceptually simpler or more appropriate tend to get obscured by the proliferation of formalism. (ii) Another novelty compared to the existing treatments is the first section dealing with phonotactics. Rather than using filters, cooccurrence restrictions, and the like, phonotactic statements are formulated as 'precyclic' phonological rules.[1] In particular, a general rule for dissimilation in initial consonant clusters is assumed to account for the lack of *tl*, *dl*, *szr*, *pv*, and *bv*. (In a subsequent article, Olsson (1993) points out that intramorphemic labial

[1] Olsson assumes a kind of lexical phonology framework in which rules applying across a morpheme boundary work at level 2, while the rules applying at all levels are postcyclic. It is unclear whether the latter subdivide into postcyclic lexical and postlexical rules; also, while precyclic rules appear to be located at level 1, it remains unstated whether level 1 as a whole is assumed to be non-cyclic (precyclic) or there are also other level 1 processes that apply cyclically.

sequences are restricted to nasal-initial (intervocalic or final) clusters anyway, so the lack of *pv*, *bv* need not be stated as part of the general dissimilation rule; that move reduces the scope, if not the generality, of dissimilation.) (iii) A three-valued rule feature convention is also proposed, such that [+rule n] triggers the application of a (minor) rule as usual (or else goes lexically unspecified for general rules), [0 rule n] blocks the application of a rule (this function is normally attributed to [–rule n]), whereas [–rule n], in a rather curious manner, triggers a process that is the 'opposite' of what rule n says. (For instance, this is used for disharmony on *híd*-type items (*híd* 'bridge' [–Harmony]) as opposed to [0 Harmony] for *-nék* as in *aludnék* 'I would sleep'.) Although Scots law and Aristotle are both cited as supporting evidence, the degree of convincingness of this solution is rather low; another addition to the arsenal of formal complexities. (iv) Rather than letting the outputs of vowel rules be subject to adjustment rules (as Vago does), Olsson proposes two alternatives to absolute neutralization. The first is feature hierarchy, the second an algorithm for pairing relations. Both devices are extraphonological, to say the least. (v) Several 'as if' rules are proposed which do not specify actual structural changes (in other words: do not have any effect on the phonetic output); rather, they say that certain segments or sequences behave 'as if' they were something else. Although one of these rules accounts quite nicely for the asymmetric behaviour of *j*, *h*, and *v* ('glides' in this analysis), the idea is, again, unorthodox in contemporary linguistics. Finally, (vi) "symbolic logic deriving from Peirce, Boole, Frege and later scholars has been a source of inspiration. In addition a more recent extralinguistic development has been influential—namely prototype theory. A kind of logical phonology has thus emerged" (8).

The book includes an Introduction, five chapters (1. Phonotactics; 2. Consonant rules; 3. Vowel rules; 4. Non-verbal inflections; 5. Conjugation), several indices including a very useful List of rules, and extensive references.

In the Introduction (8–21), a general 'aims, scope and previous treatments' section is followed by the presentation of the consonant and vowel inventories and their analysis in terms of distinctive features; the introduction is then concluded by a brief explanation of the rule formalism used in this book.

In the classification of consonants, the general trend is to reduce the number of place-of-articulation classes, accounting for predictable place differences in terms of manner features. By contrast, Olsson collapses the manner classes of stops and affricates but differentiates six places, as follows (a seventh place would be Glottal for the glide *h* but it is argued that the underlying segment is the velar fricative, deconsonantalized prevocalically into a glide by rule):

	Labial	Dental	Alveolar	Palato-alv.	Palatal	Velar
Nasals	m	n			ny	
Stops/affricates	b	d	dz	dzs	gy	g
	p	t	c	cs	ty	k
Fricatives			z	zs		
	f		sz	s		x
Liquids		l	r			
Glides	v				j	

(The rows for liquids and glides could also be collapsed but their distinct behaviour, Olsson claims, warrants keeping them apart.) Given that stops and affricates are united into a single class, one would think that the old debate concerning whether *ty* and *gy* are stops or affricates is resolved in an elegant manner. However, as will become clear, the idea of economy in the use of distinctive features is not one of the author's main concerns. Twenty different features are defined and used, at least one third of which are predictable from the others, hence redundant. One of these redundant features, given the above table, is 'gradual release' (SPE: delayed release). Accordingly, Olsson takes a stance on the affricacy issue with respect to *ty* and *gy*, claiming that they are affricates.[2]

With respect to vowels, the mid front unrounded vowel *ē* is part of the inventory, although its presence is neither crucially involved in the formal working of the system nor claimed to be underlyingly present in what are called 'seven-vowel dialects'. The panlectal treatment complicates the discussion without any gain in explanatory power; and the fact that forms are cited in their eight-vowel version is a constant source of misprints/misrepresentations (to mention just a few: *képp(ēn)* [read: *képp(en)*] '-ly', p. 76; *koncert* [read: *koncērt*] 'concert', p. 78; *kertem* [read: *kertēm*] 'my garden', p. 118; *pēhēly* [read: *pehely*] 'fluff', also: *pēlyhēk*, *pēlyhēm* [read: *pelyhek*, *pelyhem*] p. 128; *nekem* 'for me', *neked* 'for you', [read: *nekēm*, *nekēd*] p. 143; *mēlyikēt* [read: *melyiket*] 'which one', p. 183).

Turning to the distinctive features, [vocalic], [labial], [gradual release], [frication], and [distributed] are completely redundant, and [long], [high] and [back] are redundant for consonants. The definition for [consonantal] includes the sentence "Non-continuous sounds and liquids are consonantal": this would exclude fricatives; but it must be a misprint, as the chart on p. 16 correctly specifies fricatives as [+cons]. The definition for [vocalic] says "The most radical constriction in the oral cavity does not exceed the one that is characteristic for high vowels ... Vowels and liquids are vocalic": note that the definition would exclude liquids; this time, no misprint is involved as liquids are characterized as [+voc] throughout. The feature [dental] is defined here as "The tip of the tongue is articulator and may lie in front of or at the alveolar ridge. The feature can primarily be used to distinguish alveolars from dentals" (14).[3] Given the classification of consonants as presented in the table above, this feature is genuinely necessary and is put to good use in several cases in the book. Finally, the use of [long] for consonants is commented on as follows: "Hungarian long consonants may at times be classified as two consonants, at times as geminates (sic, read: as single segments characterized as [+long]). Compare this with the analysis of light—which, as Bohr pointed out, has both wave and particle properties" (17). The analogy sounds rather convincing; however, this alleged dual nature of long consonants is not exploited except on p. 175 where we read "For convenience, [+long] is used in (40) and two *t*'s in (37)". I fail to see even the convenience, let alone the necessity, of this move: it could just as well have been the other way round.

The first chapter (Phonotactics, 22–50) describes initial, medial, and final consonant clusters in terms of 'precyclic' rules (see above). With respect to initial clusters, the discussion is based on Siptár (1980), although the formalization is Olsson's own.[4] Consider the following (as an example of the formal device of 'subset specification', referred to above):

[2] In Olsson (1993) he explains that one should take the variant in the strongest position as basic. Given that the oral palatal noncontinuants are phonetically realized as affricates before a stressed vowel and as stops in various weaker positions, it follows that they are underlyingly affricates. This reasoning does not tally with Olsson's own treatment of *h*; and, more substantially, it disregards the fact that genuine affricates are never realized as stops, no matter how weak the position.

[3] In Olsson (1993, 163), the feature is redefined to include labiodentals in the [+dental] class; thus rendering the feature [distributed] even more redundant than in the present book.

[4] One notable departure from the source of data is that *szf* (as in *szfinx* 'sphinx') is not regarded as a legitimate initial cluster. This is reasonable; but part of the explanation says "it is probably homophone with *szv*" (as in *szvit* 'suite'); which is plainly false.

"(25) Initial cluster specification

$$+ \quad \underset{\left[\begin{array}{l}+\text{cont:}\\ -\text{voi}\end{array}\right]}{\overset{\text{C}}{\downarrow}} \quad \underset{\left[\begin{array}{l}-\text{son:}\\ -\text{cont}\\ -\text{voi}\end{array}\right]}{\overset{\text{C}}{\downarrow}}$$

(The first consonant in an initial cluster is voiceless if continuous, while a following obstruent is voiceless and non-continuous.)" (46).

Notice the 'implicative colon'; it defines an if-then relationship. The rule says, in effect, that the first consonant in an initial cluster may or may not be continuous; but if it is continuous, it is voiceless. Similarly, the second consonant may be either an obstruent or a sonorant; in the former case, however, it can neither be voiced nor a fricative. In conjunction with other rules, (25) correctly defines the set of initial two-member clusters as well as the first two consonants of three-member clusters.

What remains unclear throughout, however, is whether the domain of phonotactic rules is the syllable or the morpheme. (The third possibility, the word, is excluded by the statement that phonotactic rules are precyclic, i.e. they do not apply across morpheme boundaries.) From the rules themselves it would appear that the domain intended is the morpheme, and the fact that medial (intersyllabic) sequences are also covered turns this impression into almost a certainty; however, syllable-related terminology is often used. This is confusing in itself; what is even more confusing is that the terms are not clearly defined. Two hints are given: "(I use limit here as a cover term for onset and offset.)" (41) and "(Margin is Vennemann's 1988 cover term for head and coda.)" (44). To make this coherent, one could try the following interpretation: 'onset' is the first consonant in a syllable, not the usual 'any number of consonants preceding the nucleus'; 'offset' is correspondingly the last consonant, not the whole coda. 'Head' is the term used for what is usually called onset: the whole initial cluster; and 'coda', as usual, is the whole final cluster. In that case, 'limit' would stand for the two extreme consonantal positions (i.e. those next to a syllable boundary), and 'margin' would refer to the two nonnuclear syllable constituents. However, a rule called Limit sonorant sequence specification (42) says "A consonant that is separated from a morpheme-boundary by a sonorant is a preceding sonorant". (In other words, sonorant + sonorant and sonorant + obstruent are both excluded initially, and obstruent + sonorant is excluded finally.) This would imply that the second consonant from the boundary (in whichever direction) is also part of the 'limit'. No attempt is made to resolve this contradiction.

The second chapter (Consonant rules, 51–74) is where Olsson's notational innovations are at their best. In particular, Glide voicing assimilation, *x*-deletion and *x*-deconsonantalization, Liquid assimilation, Nasal assimilation, and Labiodental assimilation are all superior (simpler, more revealing) than the corresponding rules in Vago (1980), even if most of the formal simplicity is due to the undercurrent of autosegmental insights incorporated in Olsson's formalism. Voicing assimilation, on the other hand, is made simple by a clever way of getting around the problem of *v* that undergoes but does not trigger voice assimilation. Vago's solution involves [+son] for *v* and has to include a special rule for its devoicing in e.g. *szívtelen* 'heartless'. Olsson goes a step further: he analyses *v* as a glide (i.e. [–cons, –voc, +son]); this is not specifically motivated but does not bear on the issue at hand. The real difference between the two treatments is that Olsson posits a rule of Structural *v* strengthening where 'structural' refers to the fact that *v* is not supposed to change into something else (viz., an obstruent) but just behave 'as if' it were [–son] before a consonant or pause. Given this rule, the asymmetric behaviour of preconsonantal *v* (an obstruent that undergoes voice assimilation) and postconsonantal *v* (a sonorant that does not trigger voice assimilation) is explained away. Granted the possibility of 'as if' rules, one might speculate about doing the whole thing in reverse: analysing *v* as an obstruent ([+cons, −voc, −son]) and positing a rule of 'Structural *v* weakening' that specifies it as 'behaving like' a sonorant after a consonant (occurring as second member of initial clusters and failing to trigger voice assimilation). The consequences of this reversal might be far-reaching and will not be explored here; but the idea is appealing since the surface representation of *v* is, after all, a fricative, even if the least fricatival of all.

The area of coronal assimilations is another instance of radical improvement over Vago's solutions. The most important insight here, and one that was overlooked by Vago, is that *l* is never palatalized by segments other than *j* (cf. *elnyel* 'devour', *ökölnyi* 'fist-sized', *hölgy* 'lady', *kopoltyú* 'gill'). A number of complications are thereby avoided. (i) The $l \longrightarrow j$ / ___ j change (as in *éljen* 'let him live', *hol jártál* 'where have you been') can be formulated quite independently of the rest of Palatalization (and can be collapsed with Liquid Assimilation as in *balra* [rr] 'to the left', a marginal gain in overall simplicity). (ii) In Vago's framework, *l* was analysed as [−cont] on the basis of the sole evidence that it 'patterned with' *t*, *d*, and *n* with respect to palatalization. Although the literature has been divided with respect to the continuancy of *l* ever since SPE, the fact that Olsson redefined it as [+cont] gives a more consistent system in which all nonnasal sonorants and all fricatives are continuants, whereas all stops (oral as well as nasal) are noncontinuants. (iii) Vago's rule of Palatalization as applied to *l* yielded an intermediate palatal lateral /ly/ that had subsequently to be neutralized with [j]. This formal hitch is automatically avoided in the present treatment. (iv) Given that the assimilation of *n* to palatals (as in *kenje* 'let him smear', *pattantyú* 'cracker', *ötven gyerek* 'fifty children', *olyan nyakas* 'so stubborn') is independently accounted for by Nasal Assimilation, the rule of Palatalization need only refer to *t* and *d*, i.e. the dental oral stops.

This rule, in turn, can now be collapsed with Stop + Affricate Affrication (as in *két cipő* 'two shoes', *vad csók* 'wild kiss', *nagy dzsungel* 'big jungle') and with Strident Consonant Assimilation (as in *házsor* 'row of houses', *édes cica* 'sweet kitten', etc.) into a very general rule of Coronal Place Assimilation ("In a coronal sequence, an obstruent assimilates in place to a following consonant that is strident unless high [i.e. is either strident or palatal]. A strident obstruent only assimilates to a strident consonant" (74).) Similarly, Vago's *j*-Stridency Assimilation (*hoz* + *j* ⟶ [hoz:]) and *j*-Palatal Assimilation (*hány* + *j* ⟶ [haːɲ:]) are collapsed as *j*-assimilation[5] and his Stop + Fricative Affrication (as in

[5] This rule is formulated as in (i) below; an overuse of Greek letter variables is apparent here. [+cor, αhigh, αback] is intended to capture the class of non-palatal coronals; given that all coronals are [−back], a more appropriate characterization is [+cor, −high] as in (ii):

(i)
$$\begin{array}{c}\left[\begin{array}{cc}\multicolumn{2}{c}{+\text{cor}}\\ \begin{bmatrix}\alpha\text{high}\\ \alpha\text{back:}\\ +\text{strid}\end{bmatrix} & \begin{bmatrix}-\text{cons}\\ +\text{high}\end{bmatrix}\end{array}\right]\\ \downarrow \quad \searrow \\ [\qquad\qquad\qquad]\end{array}$$

(ii)
$$\begin{array}{c}\left[\begin{array}{cc}\multicolumn{2}{c}{+\text{cor}}\\ \begin{bmatrix}-\text{high:}\\ +\text{strid}\end{bmatrix} & \begin{bmatrix}-\text{cons}\\ +\text{high}\end{bmatrix}\end{array}\right]\\ \downarrow \quad \searrow \\ [\qquad\qquad\qquad]\end{array}$$

(A *j* is fully assimilated to a preceding consonant that is either strident or palatal. Note that the slanting arrow is meant to indicate full assimilation of the second segment to the first within the sequence.) A similar simplification of rule (49), Coronal place assimilation (p. 74) is also possible (hence: necessary).

ötször 'five times', *egység* 'unity') and Affricate + Fricative Affrication (as in *kilencszer* 'nine times' *ínyencség* 'delicacy') are now united as a single rule of Affrication. Thus, a set of seven different rules is reduced to a set of three.

The interaction of these rules, however, is somewhat problematic. Affrication and *j*-assimilation are characterized as Level 2 rules; whereas Coronal place assimilation is postcyclic. Now, cases like *hátsó* 'rear (adj.)' are appropriately accounted for as /t/+/ʃ/ ⟶ t + tʃ (Affrication) ⟶ [tʃ:] (CPA); however, cases like *látja* 'he sees it' would require /t/+/j/ ⟶ c + j (CPA) ⟶ [c:] (*j*-assimilation). The reader is left wondering if a postcyclic rule should be allowed to feed a Level 2 rule. Given that a 'postcyclic' rule is defined as one "applying at all levels" (21), this may simply be a case of misnomer; but it is difficult to assess whether more serious problems are not at stake here, given the general mist that surrounds these issues in the book (cf. note 1 above).

One topic that the reader misses in this chapter on consonant rules is degemination. Much later, on p. 155 (in the context of the short subjunctive *várj* 'wait!', *várd* 'wait for it!') it is casually remarked that "This rule [i.e. degemination] is trivial and will therefore not be formally stated. It consists of degemination of a long consonant when next to another consonant". In fact, however, degemination is a rather intricate issue in Hungarian—see Nádasdy (1989).

The third chapter (Vowel rules, 75–99) includes a brief summary of the vowel lengthening issue (*fa* 'tree' ~ *fát* acc.) and a more detailed, though inconclusive, discussion of vowel harmony. Apart from the three-valued rule feature convention (cf. (iii) in the first paragraph above) and a speculative section on alternatives to Vago's adjustment rules (cf. (iv)), this chapter is little more than a review of issues raised in the literature (in particular, Olsson's rules of Harmony and Rounding Harmony are neither simpler nor more revealing than Vago's corresponding rules; indeed, not even crucially different).

The fourth chapter (Non-verbal inflections, 100–144) describes the declension system and also includes short sections on adjectives, pronouns, infinitives, 'quantity words', and 'postpositions and case stems', respectively. The discussion follows the footsteps of Vago's survey of the issues that emerge; introducing minor corrections and additions where appropriate.[6] The fifth chapter (Conjugation, 145–196) discusses the verbal paradigm, introducing a number of nonformal principles like 'avoidance of homonymy' (159–163), cf.

[6] The stem types represented by *kő* 'stone', *fű* 'grass', *ló* 'horse', *szú* 'woodworm', and *lé* 'liquid' on the one hand and by *hó* 'snow' or *tó* 'lake' on the other are jointly accounted for by a complex schema which is sensitive to the voicing quality of the preceding consonant if the vowel is *ó*: thus, *ló* and the items that contain other vowels will alternate with *-öv-*, *-üv-*, *-ov-*, *-uv-* or *-ev-* whereas *hó* etc. will alternate with *-av-*. This is rather clever (if counterintuitive); however, there are several problems with this analysis. (i) There is no indication that the rule is a minor rule, applying only to inputs that are specifically marked to that effect in the lexicon. Thus it over-applies to items like *só* 'salt', *bú* 'sadness', *nő* 'woman', *tű* 'needle' etc. (Actually, the fact that the stem must be monosyllabic is not shown in the rule either; this increases the range of misapplication greatly, cf. *felhő* 'cloud', *tapló* 'tinder', etc.) (ii) No account is offered of the related but distinct group of *mű* 'work (of art)', *bő* 'loose' etc. where vowel length is retained. (iii) The verbs *lő* 'shoot', *nő* 'grow', *fő* 'cook (intr.)', *sző* 'weave', *ró* 'scribble' (177–8) are analysed the other way round: i.e. with underlying *-öv-*, *-ov-* and a rule of lengthening/deletion; no reason is given for the difference. (Also, *rí* 'cry' and *nyű* 'wear down' are assumed to undergo the same rule although these do not alternate in terms of vowel length. If, however, we represent the latter two with long underlying vowels, there is no way to safeguard *hív* 'call', *óv* 'caution (verb)' etc. from losing their *v* before a consonant or word finally.) (iv) Assuming that the minor status of the rule is acknowledged and disregarding the other complications concerning verbs and non-shortening nouns/adjectives, one crucial objection remains: the alleged bifurcation between

Carstairs (1987, 119), 'the inflectional rests principle' (168–189), based on somewhat questionable data (like ?*festēl* vs. *festesz* 'you-sg. paint', ?*áldol* vs. *áldasz* 'you-sg. bless', ?*áldtam* vs. *áldottam* 'I blessed', ?*áldni* vs. *áldani* 'to bless'), as well as a lengthy speculative section on 'the role of markedness for the conjugations' (181–196).

In the rest of this review, let us indicate a number of minor errors, oversights, and misprints that could easily be corrected in subsequent editions.

Glosses for Hungarian examples cited are mistaken in quite a few cases: *tviszt* 'twist (the dance)' is glossed as 'strife' on p. 23; *újsága* 'its newness; the fact that it is new' appears as 'his news' on p. 111; *szomszédoméi* 'those of my neighbour' is translated as 'his neighbour's' on p. 114; *hó* 'snow' is glossed as 'grass' and *szó* 'word' as 'maggot' on p. 119; *több* 'more; several' and *sok* 'many; much' appear as 'many' and 'much', respectively, on p. 138; and the case stem/postposition paradigms presented on pp. 142–3 are mis-glossed as follows: sublative *rám*, *rád*, etc. should be 'onto' (me, you, etc.) rather than 'into'; delative *rólam*, etc. should be 'away from (top of)' or something like that (also possibly 'about', a secondary but more frequently used meaning) but not 'out of'; causal-final *értem*, etc. should be 'for (the sake of)', rather than 'because of', the latter being *miattam*, etc.; and *mögém*, etc. should be 'to behind' rather than 'from behind' which is *mögülem*, etc.[7]

On p. 107, Olsson concludes that VT (Vowel Truncation) should be restricted such that it does not apply to case suffixes; but this restriction is not added either on p. 106 where the rule is first stated under (10) or (more crucially) on p. 205, in the List of Rules, where the rule of VT is mistakenly numbered (15). Also on p. 107, a sentence begins "As shown in (11)," but a corresponding item (11) is not given anywhere.—There is of course an item (11) in the next section but not the one referred to in the text here.—On p. 118, it is claimed that lowering stems prevent *j*-insertion. This tends to be true but there are too many counterexamples for the claim to be made in such absolute terms; cf. *tál* 'dish': accusative *tálat* (lowering) but 'his dish': *tálja* (*j*-insertion); similarly: *kád* 'tub': *kádat* but *kádja*; *út* 'road': *utat* but *útja*; *föld* 'land': *földet* but *földje*; *sült* 'roast (noun)': *sültet* but *sültje*; *híd* 'bridge': *hidat* but *hídja*, etc. Conversely, there are a number of non-lowering stems without *j*-insertion, cf. *kár* 'damage': *kárt* but *kára*; *sör* 'beer': *sört* but *söre*; *szám* 'number': *számot* but *száma*; *gyök* 'root': *gyököt* but *gyöke*.—In (27b) *kert* 'garden' is cited as an example of lowering stem but the correct 8-vowel form is *kertēm* 'my garden', hence *kert* is regular.

Rule (50) on p. 133 introduces the syntactic features [±N], [±V]. Thus V in a rule is systematically ambiguous between 'vowel' and 'verb'; an unfortunate result. On p. 135 the representative form *iparost* 'craftsman-acc.' is derived by what is known as the 'Duke of York gambit', i.e. epenthesis followed by syncopation (*iparos-t* ⟶ *iparosot* ⟶ *iparost*). Even if this solution is directly borrowed from Vago (1980, 94–5), it cannot be regarded

---

voiceless consonant + *ó* and the rest of cases does not work. The word *jó* 'good' consists of a voiced consonant followed by *ó*, yet patterns with *hó* and *tó* (cf. *javak* 'goods', *java* 'its best/largest part', etc.). It is true that the accusative is *jót* rather than *javat*, but this is exactly paralleled by *szó* 'word' that Olsson explicitly includes in the *hó* class, and anyway the predicted **jovat* is definitely out. Thus, the *lovat* 'horse-acc.' vs. *tavat* 'lake-acc.' difference cannot be attributed to the voiced vs. voiceless consonant that precedes the *ó*.

[7] Miscellaneous misprints include 'esparagus' (read: 'asparagus'), p. 23; 'plural allomorph for possessions as A' (read: '. . . as Ai'), p. 112, lines 1-2; 'Mini-Length' and 'Mini-Low' (read: 'Min-Length' and 'Min-Low', respectively), p. 129 (44); '*mos* + *sz* ⟶ **moss* [mɔʃ:]' (read: '*mos* + *sz* ⟶ **mossz* [mɔs:]'), p. 152, last line; '*lássak*' (read: '*lássam*'), p. 158; 'van der Hulst 1983' (read: 'van der Hulst 1985'), p. 160; 'mēgbusszol', 'mēgbusszolt' (read: 'mēgbosszul', 'mēgbosszult'), p. 172; and 'nyugodt' (read: 'nyugodott'), p. 179.

as insightful. This could have been another instance of simplification over Vago's system: Olsson's rich arsenal of notational devices surely would have made it possible to do it some other way.

On p. 137 it is claimed that infinitives inflected for person (e.g. *várnom* 'for me to wait') are "not used with prepositions, as this class is very restricted in the language". Very restricted indeed, to the point of non-existence.—The third person infinitives (e.g. *várnia* 'for him to wait', *várniuk* 'for them to wait' are claimed to involve deletion of *i* as in other persons, followed by *j*-epenthesis as in the nominal paradigm, thus *várn-om*, *várn-od*, *várn-ja*, cf. *bárd-om*, *bárd-od*, *bárd-ja* 'my/your/ his axe'. "The spelling should then just be traditional rather than phonetic, reflecting an assumed connection with the suffix *ni*" (138). There are two reasons why this clever suggestion is untenable. First, a form like *várnia* is trisyllabic (witness the Y/N question test: *Kell vár↑ni↓a?* 'Does he have to wait?', **Kell ↑vár↓nia?*); second, the analysis *i*=[j] would predict *[vaːrɲa] by Nasal Place Assimilation + *j*-assimilation.

On p. 143 it is claimed that *emögött* 'behind this' is pronounced [ɛmːøgøtː] "(the spelling is not phonetic for postpositions)". It is also pointed out that *-ja/je* (as in *alája* '(to) below it', *melléje* '(to) beside it', *közéje* '(to) between (pairs of) it', *mögéje* '(to) behind it', etc.) "occurs only in substandard speech" (144). To this particular native speaker of Hungarian (the reviewer), at least, **emmögött* sounds a lot more substandard than *mögéje*. Similarly, it is correctly suggested on p. 177 that the traditional first-person ending of *-ik* verbs (as in *mászom* 'I climb') "seems to be on its way out"; but the example cited ("*k* is used for e.g. *lakik* 'inhabit'") is not very aptly chosen: *lakok* 'I live (swhere)' strikes me as rather more substandard than *mászok* 'I climb' (the latter being labelled as such by Olsson).

In the section entitled 'The underlying pexes' (163–4) the list of pexes (= person-number complexes) includes *lAk* and *tOk* but the explanatory note says "The forms are given without the epenthetic vowels that may follow a verb stem; remaining epenthetic vowels are not indicated either" (163); given that the *A* of *lAk* and the *O* of *tOk* turn out to be regular epenthetic vowels, the list should just include *lk* and *tk*, respectively.

In (28) on pp. 164–5 the paradigm of *áld* 'bless' indicates extensive free variation between *áldasz/áldsz*, *áldalak/áldlak*, *áldotok/áldtok*, *áldanak/áldnak*, and similarly throughout the past (*áldottam/áldtam*) and conditional (*áldanék/áldnék*) paradigms. This variation (pace Vago 1980, 59) is simply nonexistent (the shorter forms are unacceptable). Furthermore, on p. 168, another variant for *áldasz/áldsz* is mentioned: *áldol* (which is just as nonexistent). Trying to explain why **áldol* is not **áldal* instead, Olsson suggests that forms like *áldol* "may be compared with the superessive plural of lowering stems like *ház*, where the vowel before the superessive suffix is also unexpectedly mid—*házakon*". The reader wonders why it is not either the singular of lowering stems (*házon* 'on a house') or the plural of *any* noun stem (*gázokon* 'on gases') is cited; *-on* is resistent to lowering, whether triggered by stem type (cf. *házat* 'house-acc.') or by the fact that there is an intervening suffix (cf. *gázokat* 'gases-acc.'); the choice of *házakon* 'on houses' wrongly suggests that the expected (but not attested) lowering in such forms would be due to stem type rather than plurality.

Despite the above objections, most of which concern matters of detail, it remains true that Olsson's dissertation is a major signpost along the long and winding road to exploring Hungarian phonology and morphology; it represents a major step forward in the right direction, even if couched in a formalism that sometimes makes it difficult to appreciate the advantages of this treatment over previous attempts.

Péter Siptár

References

Abondolo, D. 1988. Hungarian Inflectional Morphology. Akadémiai Kiadó, Budapest.

Carstairs, A. 1987. Allomorphy in Inflexion. Croom Helm, London.

Hall, R.A. 1944. Hungarian Grammar. Language Monograph No. 21. Linguistic Society of America, Baltimore.

Hulst, H. van der. 1985. Vowel harmony in Hungarian: a comparison of segmental and autosegmental analyses. In: Hulst, H. van der–Smith, N. (eds): Advances in Nonlinear Phonology, 267–303. Foris, Dordrecht.

Kornai, A. 1986. On Hungarian Morphology. Linguistica, Series A, 14. Linguistics Institute of the Hungarian Academy of Sciences, Budapest.

Nádasdy, Á. 1989. The exact domain of consonant degemination in Hungarian. In: Szende, T. (ed.): Proceedings of the Speech Research '89 International Conference, 104–107. Linguistics Institute of the Hungarian Academy of Sciences, Budapest.

Olsson, M. 1993. The basic Hungarian allophone system: Structure and rules. In: Lund University Department of Linguistics, Working Papers 40: 157–84.

Siptár, P. 1980. A note on inital clusters in English and Hungarian. In: Acta Linguistica Hungarica 30: 327–43.

Siptár, P. 1984. Review of Vago 1980. In: Acta Linguistica Hungarica 34: 138–45.

Vago, R.M. 1980. The Sound Pattern of Hungarian. Georgetown University Press, Washington.

Vennemann, Th. 1988. Preference Laws for Syllable Structure and the Explanation of Sound Change. Mouton–de Gruyter, Amsterdam.

Julián Méndez Dosuna and Carmen Pensado (eds): Naturalists at Krems. Papers from the Workshop on Natural Phonology and Natural Morphology. Ediciones Universidad de Salamanca, Salamanca 1990, 127 pp.

0. In this review of a collection of papers, I will try to browse through the main ideas presented in the ten papers focusing my attention on the pieces discussing phonological issues. The editors of the volume are professors at the University of Salamanca. They have collected papers delivered at the Workshop on Natural Phonology and Natural Morphology, which was held as part of the Sixth International Phonology Meeting and Third International Morphology Meeting in Krems on 1–7 July 1988.

The publication looks very attractive at first glance, but it contains a surprising number of annoying errors. The sloppy IPA symbols may be excused by the insufficiency of the text formatter used by the editors, it is more unlucky that they oftẹn miss to supply the symbols they could in no way produce by the software. On page 47, we find a transcription of the Polish sentence *Pan, nie Pani!* as [pʰan, e pʰa i]. (The comma is also unusual in an IPA transcription.) On the next page it is again the task of the reader to guess from the context that a retroflex approximant is missing from between the two slants, "/ /" is all we can see. A page later the transcription of English */bomb/* appears as [b m]. At the middle of the same page a Polish example is misspelt *reka* ['rȩŋka] instead of *ręka* ['rɛŋka]. The chart on pages 77 and 78 was broken somewhat awkwardly, the heading does not appear on top of the continuation of the chart leaving it hard to interpret. Hyphens are used to represent the em-dash, but this is done inconsistently: in certain cases we find one hyphen with spaces on both sides, in others two hyphens with space only on one side, in still others without a space on either side. The numbers referring to footnotes sometimes find themselves before the punctuation mark, sometimes in their usual place, after it. Instead of a reference (or was an example intended?), we find cf. 000 on page 99. The spelling of the texts was not checked, it abounds in misspelt words: e.g., *paradygmatic, synchretism* (29); *developping, non_generalization, non_application* (84); *dissyllabic, to* instead of *too* (85). But rather than lingering on the appearance let us look at the contents.

1. After a short preface by the editors, in which the authors of the volume are listed and acknowledgments made, the first paper (they are arranged in alphabetical order of the authors' names) is by Peter Auer, titled "A Note on Prososdy in Natural Phonology". The paper deals with an issue not treated before in the framework of Natural Phonology (NP). Auer does not agree with the extreme view that processes that have their basis in the articulation system are the only candidates for being a natural explanation for phonological phenomena.

A short excursus has to be made here on the status of rules in natural phonological and morphological theory. Dressler (1984) defines rules as morphonological (i.e. morphophonological) and allomorphic morphological rules, which are not productive and unexplainable synchronically. His examples for this type of alternation are umlaut phenomena in German and English and the Great Vowel Shift of the latter. These he equates with Kiparsky's lexical rules. As opposed to this, Dressler calls post-lexical rules processes, which he divides into two groups: dissimilatory (fortition, foregrounding/clarification) and assimilatory (lenition, backgrounding/obscuration).

In the terminology of Natural Theory, it is the set of processes then that Auer claims to be containing processes conforming to general principles of cognitive processing. This claim is backed by the fact that the aim of Natural Theory is explaining morphological and syntactical as well as phonological phenomena by natural processes, which then obviously cannot be limited to the speaker's articulatory and the listener's perceptual ease. To support his view, he cites an example from Schein–Steriade (1986): there is a rule in Tigrinya which fricativizes postvocalic [+back] stops. The rule is blocked if the stop is part of a geminate cluster, thus /qätälä-kka/ 'kill-(perf.)-2, masc.'='s/he killed you' will not become *[qätäläxka], but not if a morpheme boundary intervenes: /barak-ka/ 'bless-(perf.)-2sg. masc.'='you blessed' surfaces as [baräxka]. Auer says that the fricativization process, which facilitates the articulation in heterogeneous consonant clusters, would reduce pronounceability if it were applied to geminates. There is no real reason though why we should claim [kk] to be more readily pronounceable than [xk]. But this is not Auer's main point yet, his problem is if it is the ease of articulation that blocks the application of the process for geminates, why does it still work if there is an intervening boundary. He proposes that the solution is a semantic (or morphological) teleology: morphemes must not be endangered by phonological phenomena, this is a natural principle maintaining the ease of cognition. Thus what saves the geminate not divided by a boundary is that for semantic/morphological reasons the form of the morpheme is not distortable. We may wonder why the same principle is relaxed in the case of the morpheme /barak-/. In a footnote Auer compares the explanation present-day generative phonology and NP give to the problem. Schein–Steriade say that geminates resist change because in their representation the same segmental material is associated with two skeletal points, thus forming a relatively stable configuration. This—Auer says—is no more than a formal description of what happens. But his explanation does not explain either why the form of the stem /barak-/ might be altered, in fact, it is much more plausible to think of monomorphemic geminates being the same material associated to two slots, while virtual geminates (with an intervening boundary) being two feature bundles hanging from two slots, as Schein–Steriade probably do, in which case neither of the slots looks by itself as part of a geminate, therefore restrictions applying to geminates have no effect on them.

At this point, Auer goes on to examine prosody in a natural framework. Here again, the needs of articulation and perception will not suffice as an explanation, rhythm will find its explanation in the needs of cognition. Language, like other forms of human behaviour, is rhythmically organized, and in this organization, there is always some kind of isochrony: an approximately even temporal spacing. In most languages of the world, we find one of two systems of isochrony, one based on the syllable, the other on the foot. Auer mentions a third, too, based on the mora, but the features correlated with it are similar to those of the syllable-based type, so he does not deal with it any longer. Languages with rhythm based on the syllable (syllable-timed rhythm; iso-syllabic) are more phonologically oriented, while those with their rhythm based on the foot (stress-timed rhythm; iso-accentual) are more grammatically (semantically) oriented. Auer considers phonological phenomena which may be connected with the two timing systems. He refers to the only previous paper concentrating on the problem: Donegan–Stampe (1983). That paper examines the Munda (iso-syllabic) and Mon-Khmer (iso-accentual) language families, establishing the following correlations:

- syllable rhythm (syllable timing)—case system—suffixing—left-bound ('falling') word accent—falling phrasal accent—operator-last syntax (i.e. modifier before modified);
- word rhythm (stress timing)—no case system—prefixing—right-bound ('rising') word accent—rising phrasal accent—operator-first syntax.

Auer does not agree that these correlations could be generalized and cites European languages that prove to be counterexamples. In the following discussion, he restricts himself to these languages and to phonological phenomena, using English as a system having stress-timed rhythm and Italian as one with syllable-timed rhythm. Consequences of stress-timed rhythm are the reduction of unstressed syllables and thus the loss of contrasts in this position; a high degree of cliticization, the loss of the syllabic nuclei in clitics resulting in the emergence of various complex consonant clusters, which in turn form fertile ground for assimilatory and dissimilatory processes; because of the greater importance of the foot than that of the syllable the loss of the significance of syllable boundaries, the ease of resyllabification and ambisyllabicity of single intervocalic consonants following short vowels. The coherence of the foot is maintained at the expense of clear-cut syllable boundaries by lenition processes like flapping and fricativization. In the absence of clear-cut syllable division, true geminates also fail to occur. Auer says that "in an iso-accentual language, clusters of homorganic consonants are immediately reduced to single consonants" (18). The wording is misleading: what he probably thinks of are not homorganic consonant clusters but geminates, since the claim is easily falsifiable for other homorganic clusters where the first segment is a nasal ([mp, ŋk]), a fricative ([st]) or a liquid ([lt, rt]). Foot-level compensatory lengthening is also characteristic of stress-timed languages: e.g., OE /nama/ → ME /naːm/ 'name'. The last phenomenon Auer mentions is iambic reversal and deaccentuation, which maintains the equilibrium in the length of feet. One of his examples is wrong: he cites the alleged iambic reversal in *síxteen* ~ *she was ónly sixtéen,* but the place of the lexical stress is on the second syllable of the word (*sixtéen*), from where it may move to the first in a phrase like *síxteen gírls.* (If it were on the first syllable, there would be no reason for the reversal since one unstressed syllable would still separate the two stressed ones in **she was ónly síxteen.*)

In contrast with this, in Italian, which exemplifies syllable-timed rhythm, Auer finds the opposite of the phenomena above. There is indeed no such radical reduction in unstressed syllables like in English, and although vowels in stressed syllables do lengthen phonetically, this length is not phonemic, so we cannot talk about reduction in unstressed position. The uniformity of syllables is also served by strong phonotactic constraints, like the frequent occurrence of open syllables and the fact that only these are permitted in word final position. This is in fact a strong constraint, it is not clear, however, how it contributes to the uniformity of syllables, and if it did, how we should account for raddoppiamento sintattico, which makes stressed open syllables closed. This process can in no way be phonetic in traditional terms, although it does show some similarity to stressed vowel lengthening, but then the two processes together seem to work on foregrounding the stressed syllable, which, as Auer wants to prove, is not characteristic of languages with syllable-timed rhythm. In Italian, the significance of syllable boundaries is emphasized by the great number of open syllables, the maximization of onsets and the fact that fortition occurs in onsets, while lenition in codas. The primary nature of syllable unity also excludes resyllabification and ambisyllabicity, but predicts the existence of geminates. Finally, mora compensation is limited to the syllable: syllables with a long vowel are open, those with short are closed. A possible counterexample turns up at this point again: in Hungarian, which seems to be a syllable-timed language like Italian, the loss of word final short vowels was compensated for within the foot, similarly to the Old English–Middle English change, /utu/ became /uːt/ 'road'. Against such objections Auer defends himself by saying that his classification shows tendencies only, about individual languages we can only say that they are **rather** like this type or **rather** like that type, but there exists no "perfect" language.

This might be taken as the greatest flaw of Natural Theory: it talks about tendencies, therefore its statements are hardly falsifiable and as such not conforming to the strict criteria of scientificness.

2. The second paper, "Conversion as morphological metaphor", is written by Grazia Crocco-Galeas. As the title suggests, morphological conversion, the connection between words of the same form, with similar meaning but belonging to different word classes are examined. First, the author browses through previous theories about the phenomenon, which can be divided in two main groups: according to the first of which, we are dealing with one lexical unit here that somehow belongs to more than one word classes, while the other group of opinions claims that we have real derivation here with an affix that happens to have no phonological shape. Crocco-Galeas proposes a rather different solution, he says conversion might be best compared to stylistic metaphor, the difference being that while the latter is paradigmatically paradoxical, but perfectly fine syntagmatically, conversion as a morphological metaphor may cause surprise syntagmatically, paradigmatically it is identifiable. Finally, he addresses the problem that although conversion seems to be quite unnatural a phenomenon, it is not at all uncommon compared to other, more natural ways of derivation. According to the principle of the economy of signantia, conversion does have a degree of naturalness, since one signans belongs to more signantia. Thus it may often occur especially in languages of an isolating tendency.

3. Wolfgang U. Dressler also discusses a morphological issue in his paper titled "Sketching Submorphemes within Natural Morphology". He analyses and classifies parts of words without a proper meaning (not morphemes proper) which can nevertheless be isolated within the words. With an especially rich selection of references he provides a useful guide to those interested in the topic.

4. In her paper, "A Natural Model of Acquisition of Second Language Phonology and the Notion of Relative Markedness", Katarzyna Dziubalska-Kolaczyk gives an account of her research on the acquisition of English phonology by Polish students of the language, making predictions of the degree of difficulty the acquisition of certain phonological phenomena presents to the learners based on hypotheses of Natural Theory, then introducing her findings that usually confirm the predictions. According to the theory, children are born equipped with universal phonological process types that they later on, acquiring their native language, suppress, limit and put in order. This process is subconscious, but later on the same techniques are also used in a more controlled manner in what we call learning. Acquiring the phonology of a second language means unsuppressing (relaxing the suppression of) processes suppressed in the course of native language acquisition, making limitations in new ways and reordering processes. Dziubalska-Kolaczyk found this to be much more successfully done by learners in a formal setting (language course accompanied by a teacher) than by natural setting learners, who receive little or no formal instruction.

Dziubalska-Kołaczyk examines the acquisition by Polish students of four processes that exemplify differences between Polish and English: (1) aspiration, (2) the devoicing of word final obstruents, (3) the replacement of English dental fricatives ([θ, ð]) and the retroflex approximant ([ɹ]) and (4) the problems students have with the morpheme final velar nasal ([ŋ]).

She claims the aspiration of syllable initial voiceless stops in English to be relatively unmarked as compared to Polish, where such aspiration only optionally marks emphasis. The problem of markedness might not be so simple in this case. Some accounts (e.g., Harris 1994, 133–8) show that in some languages (like most dialects of English) there is aspiration in this position, but we do not find regressive voicing, whereas in others regressive voicing does work and aspiration is absent (Hungarian could serve as an example), there is little reason to claim either type to be more unmarked than the other. (It is interesting to note that word final devoicing is predicted to occur in the second type—without aspiration, with regressive voicing—by Brockhaus (1990), but it tends to be found in languages of the other type.) It is true nevertheless that in this classification Polish belongs to the first type, so it is surprising (i.e. marked) that there is no general aspiration. Dziubalska-Kołaczyk is confirmed by her research results: since English is more unmarked, applying and not suppressing the allegedly universal process, it is relatively easy for Poles, speakers of the marked language, to acquire the process.

The universal unmarkedness of word final devoicing is again disputable (see above), but if we were to agree with Dziubalska-Kołaczyk, we would expect speakers of the unmarked (devoicing) language, Polish to have difficulties, since this universal process is suppressed in English. This claim of the author is not true for most English dialects, including both of the standard dialects usually taught to foreigners, where word final obstruents are devoiced (cf. Jones 1956, 120; Gimson 1989, 153 and 180). The devoicing is phonetic, the obstruents still behave as voiced lengthening the preceding vowel. The research proves that Polish speakers do have difficulties: according to Dziubalska-Kołaczyk, this is because they have to suppress a natural process.

As opposed to the previous processes, the replacement of English [θ] and [ð] by [t/s/f] and [d/z/v] respectively is a context-free foregrounding process, in which the marked English pair becomes unmarked. The changing of [ɹ] to [r] is of the same type, but here, she says, two related processes are at work: [+retroflex] $\rightarrow$ [−retroflex] and [+approximant] $\rightarrow \begin{bmatrix} \text{+vibrant} \\ \text{+trill} \end{bmatrix}$ yielding the more unmarked Polish trill. Dziubalska-Kołaczyk argues that the two concomitant processes producing [r] out of [ɹ] are easier to suppress that the three competing ones replacing [θ/ð], which means that Polish learners ought to acquire [ɹ] more readily than the fricatives. This conclusion is borne out by her data.

We may have two objections in connection with this part of the paper: (1) what reasons do we have to believe [ɹ] to be more marked than [r] (this is far from being clear in the case of the alternatives of replacing [θ/ð], too), and (2) is it justifiable within a theory which declares itself natural to make conclusions based on the number of rules/processes that another rather arbitrary system manipulating distinctive features uses to formalize a change.

The last contrasting phenomenon treated is English morpheme final [ŋ], the appearance of which is the result of two universal processes: the assimilation of a nasal in place to the following stop and the deletion of a stop following a homorganic nasal. It is hard to believe that the second process is universal for fortis stops, and even for lenis stops the loss of non-coronals seems to be more natural. According to Dziubalska-Kołaczyk, the two processes are constrained in English in the following way: the nasal assimilates in place within the morpheme and morpheme final labial and velar lenis stops are deleted. By morpheme final position, she probably means analytical boundary (#), since there is no deletion before non-analytical boundaries: *singer* [sɪŋɡ̸#ə] vs. *iambic* [aɪæmb+ɪk] and *diphthongize* [dɪfθɒŋg+aɪz]. Nasal assimilation works in Polish as well, the stop deleting process on the other hand is suppressed, therefore obstruent devoicing gets activated. The solution of English seems to be more unmarked, the universal stop deleting process works in it, while it is suppressed in Polish. But, says Dziubalska-Kołaczyk, this is compensated for by another universal process, final obstruent devoicing. Since universal processes are not universally ordered, we cannot decide which language's solution is more unmarked. This implies that if we take the devoicing process to apply first, Polish is unmarked, if the stop deleting process gets priority, English should be preferred. Here the author's original description of the stop deletion process gets her into trouble: if the universal process deletes any stop following a nasal then we may wonder why it is not Polish that follows the less unmarked route suppressing the universal stop deletion process even if only after devoicing has done its job. The data confirm the view that the level of markedness is undecidable, some students did acquire the English ordering of the processes, others did not. Dziubalska-Kołaczyk's final conclusion is that "it is easier for the learner to 'unsuppress' a universal process than to suppress a process that eliminated a universal one" (50).

5. Rüdiger Harnisch examines irregular verbs, nouns and adjectives in a German dialect of South Thuringia in his paper titled "Morphologische Irregularität—Gebrauchshäufigkeit—psychische Nähe. Ein Zusammenhang im empirischen Befund und in seiner theoretischen Tragweite". He aims at the conclusion that forms within the same semantic field have a tendency, on the one hand, towards a high degree of distinctivity and, on the other, towards easy access. Both distinctivity and availability are secured through the principle of lexical representation, and both are supported by the principle of shortness, which can serve as an additional marker to heighten distinctivity, or it can in itself facilitate availability.

6. "Natürlichkeit vs. Frequenz. Überlegungen zu betonungsabhängiger Allomorphie in der Deklination russischer Substantive" is the title of Gerd Hentschel's paper, in which he, in trying to establish criteria for morphological naturalness, investigates the implications of the stress patterns of certain Russian nouns with mobile stems. After a consideration of token frequency and the spreading-information theory of semantic processing, he asserts that the distinction of both number and case by stress patterns is syntactically and semantically motivated in Russian and correlates with the economy of perception, which plays no, or a very limited role in the case of less frequently used words, since these are much less networked semantically. Hentschel concludes that, as the example of Russian nouns shows, the notion of naturalness should not be reduced to unmarkedness, but rather it should seriously consider, for example, the significance of token frequency as well.

7. Marianne Kilani-Schoch describing the historical development of French clitics makes generalizations about their phonological structure and size. The first chart (on page 77) is rather confusing, we find clitics in the standard orthography and in IPA transcription, the reasons for the choice between the two is unclear to me, and the chart does not contain any reference to the syntactical and/or semantical functions the clitics have. The tendency is characteristic of the whole paper: the reader has to be acquainted with the proclitics of the language discussed for there is little information on that in the text, which looks as if it were written for people quite familiar with the topic anyway. There is apparently no reason for using the symbol "σ" to represent the phoneme /s/, which, together with /t/, was deleted in word final position between the 9th and 11th century (footnote 4 on page 78).

Kilani-Schoch makes the observation that Middle French aims at having an open syllabification: prevocalically the final vowel, preconsonantally the final consonant of the proclitic gets deleted. Phonologically the situation may be considered natural, but all the less so morphologically: third person verbal clitics do not refer to number, while third person nominal clitics do not indicate gender.

After the 17th century, in the period labelled Modern French the system of clitics was reorganized. Before verbs the prevocalic–preconsonantal allomorphy collapses, but it is retained, except for forms ending in *-e*, in the nominal paradigm. In français avancé the system undergoes further changes. The mixing of standardly spelt and phonologically transcribed forms is disturbing again, for example, we find the following in the first row of the first column: *j', [ʃ]*. It would have been more explicit to transcribe the first allomorph as well, since the second is also spelt by the standard orthography as *j'*, leaving us uncertain whether we have two allomorphs here, or the same morph spelt in two ways. The two data in the last row of the verbal paradigm (*ça* and *c*) are rather obscure. They never turned up in previous charts, what is more, the second one is neither orthographical, nor an appropriate IPA transcription symbol.

At the end of her survey, Kilani-Schoch summarizes the development of French clitics as an attrition of polyphonemic forms yielding, especially in unmarked cases, monophonemic clitics. The reason is perfectly natural, the decrease of semantic content. In any case, with this change French drifts from a typologically analytical language to a synthetic one, though there are changes pointing towards both the agglutinating: elimination and regularization of old inflexions (e.g., *résoudre* > *solutionner* 'to solve', [bõzɔm] > [bɔnɔm] 'chaps'); and the inflecting type: for example, the plural prefix *z-* after determiners (*quatre z-enfants* 'four children') and on adjectives (*des avions à réaction z-américains* 'American jet planes'), or the obligatory verbal prefix *il* following definite NP subjects (*le bébé, il pleure* 'the baby cries'), which may be analysed as an agreement marker between the topic and the comment.

8. The title of the editors', Julián V. Méndez-Dosuna and Carmen Pensado's paper is "How Unnatural is Spanish *Víctor* → *Vict-ít-or*? Infixed diminutives in Spanish". One of their charts also presents some difficulty to the reader, columns are shifted and mixed on page 91, and many of the glosses are missing. The only help to guess that the name *Óscar* has penultimate stress in this language is that words are grouped according to the place of stress in them. The wording of footnote 7 (92) is somewhat prejudiced, the stylistic label **substandard** is opposed to **correct.**

The paper is written on the morphology of diminutives in Spanish. In certain cases, like the one in the title, there seems to be infixation of the diminutive morpheme, which is an unnatural process making many try to prove that what we encounter here is not infixation. The common core of explanations of this type was that by treating the stem final element as some kind of class forming suffix the diminutive can be analysed as a suffix as well, preceding the final class forming suffix. It has also been suggested that some Spanish dialects delete word final consonants thus neutralizing the difference between *pérro* and *Víctor* type words and between the diminutive formation of the two types. The authors list convincing evidence against all these proposals. They claim that the diminutive affix is indeed an infix, the naturalness of which is explained by this being the only source of recoverability. An impressive number of evidence can be brought up in favour of the claim that diminutive affixation is not a lexical process: it is very productive, semantically transparent, does not change syntactical category and does not trigger the typical monophthongizing and velar softening rules of Spanish lexical phonology (e.g., *viéjo* 'old (man)' ~ *vej-éz* 'old age' vs. *viej-íto* 'little old man'; *plásti*[k]*o* 'plastic' ~ *plasti*[θ]*idád* 'plasticity' vs. *plasti*[k]*íto* 'plastic-dim.'). If diminutive forms are not stored in the lexicon (and they do not seem to be), then the listener has to reconstruct the stem. It is then to this reconstruction that the speaker provides some help by applying an unusual process (infixation), when the stem itself is also unusual: instead of the regular -$\acute{V}C_1\#$ stress pattern the words in question have -$\acute{V}C_0VC\#$. It is not only non-analytical affixation fading morpheme boundaries or the use of suppletive stems that diminish transparency so much appreciated by natural morphology, but also if there is no one-to-one association between derived forms and their possible stems. The somewhat unnatural process of infixation helps the fulfilment of this requirement in Spanish.

9. Geoffrey S. Nathan writes about the possible types of stops in a paper titled "On the Natural Phonology of Voicing". It begins with the introduction of an earlier paper on the topic, Keating (1984), according to which phonological theory needs two kinds of features: phonological and phonetic, to explain how two phonetically identical sounds can represent phonologically different categories in different languages. The features [voiced] and [voiceless] are phonological, whereas phonetic features include {voiced}, {voiceless unaspirated} and {voiceless aspirated} (the first type conventionally enclosed in square brackets, the second in braces). A diagram (108) shows the alternative realizations of the two phonological features: [voiced] is associated with {voicing lead}, yielding a fully voiced segment, [voiceless] with {short lag} and {long lag}, giving a voiceless unaspirated and aspirated segment respectively. The text below the diagram is interpretable as intended only if the diagram is mirrored. Referring to Lisker–Abramson (1964), Nathan writes that the situation is more complicated in English, where {voiceless unaspirated} may be the realization of both the phonological features [voiced] in word initial (and we should add word final) position and [voiceless] after /s/. In the latter case, however, we have no reason to suppose that [t] is the allophone of a voiceless phoneme, unless we base our opinion on orthography (or the traditional phonemic transcription, which probably also follows spelling in this case). This segment does not alternate with an aspirated allophone since if a strong boundary separates /s/ and the stop in affixation (e.g., *mis#trust*) the latter will still be aspirated, if the process takes place in the lexicon (e.g., *mistake*) the components are not normally said to be alternating with a free morpheme. Therefore it is perhaps simpler to suggest that the second segment of a word like, say, *stop* is /d/ phonologically, [t] phonetically.

Nathan considers vowels the fundamental elements of the phonic representation of language, and argues that the nature of vowels defines the kind of consonants possible. Vowels are periodic sound. Consonants then, which are opposed to vowels, are not periodic (fricatives), not sound, i.e. silence (stops), or a combination of the two (affricates). (Sonorants are transitional between vowels and consonants.) Citing physiological experiments, the author claims that the phoneme system easiest to pronounce, and therefore most unmarked, contains voiceless stops in word initial and final, and voiced ones in intervocalic position. (Since in this case there will be no voice opposition the phonological features [voiced]~[voiceless] have no role, the phonetical features, on the other hand, do.) That the sound systems of natural languages are not so simple, Nathan seeks to explain by the conflicting demands of production and perception: it goes against the requirements of the ease of articulation that the phonological making up of lexical elements needs an adequate number (two or three dozen) of phonemes which must be distinctive enough. Vowels are (on the one hand) opposed by voiceless stops. Therefore the unmarked case is if the stop is voiceless in every position, especially in intervocalic position, as this strengthens the contrast and by that means helps perception. A stop is even more distinguished from its surroundings (and from simple voiceless stops) if the vocal cords remain silent for some time after the period of their closure, this is aspiration, a foregrounding process. Another way of contrasting stops with pure silence is not to stop vocal cord vibration during their closed period. This is easier to do than not between vowels, but it is a marked, therefore foregrounding process word initially and finally. Experiments provide evidence for this hypothesis: children acquire intervocalic voiced stops much earlier than those in word initial or final position.

Nathan concludes that individual languages can make their choice of three series of stops. The simplest one chooses the system involving a synchronous closure of the oral and nasal tracts and suspension of voicing, thus having a {voiceless unaspirated} series. To supplement this, languages may pick one of two fortition processes, contrasting their unmarked series with either the {voiced} or the {voiceless} (he evidently means voiceless aspirated) set. Still other languages may opt for contrasting all three series. It is not totally clear why the requirements of perception have the upper hand. If we have a phonetically {voiceless unaspirated} series (as shown by the braces) then those will be voiceless intervocalically, too, despite the fact that this is more difficult for the articulators. The merging of the so far different levels of phonology and phonetics is perhaps justifiable, but certainly unclarified.

Finally, the author exposes an experiment carried out on the changes of stop series in native speakers of Spanish learning English. In Spanish voicing is the source of contrast, while in English it is aspiration. The results indicate that the change is significant only for the velar stops, it is only for them that learners start applying the English way of contrasting by aspiration. Nathan explains this by supposing that it is more difficult to retain the vibration of the vocal cords if the place of closure is near to them. He may indeed be supported by the fact that [g] is a "weak" segment often undergoing lenition processes like, for example, g → ɣ or g → h, more often than the other voiced stops. Unfortunately, the change is either context free as in, for example, Czech, or happens more often intervocalically and word finally than where we expect it, where it is more difficult to pronounce, word initially. Another lesson the author draws is that with the acquisition of the new, English-type fortition process, aspiration, the old one, voicing, is lost.

10. In their paper, "The Acquisition of Affricates in Viennese German", Chris Schaner and Livia Tonelli examine how Viennese children with German as their native language acquire the affricates [t͡s] and [p͡f] and the cluster [ks]. After a detailed presentation of their experiments and summarizing their results in five charts, the authors conclude the following: (1) the acquisition of affricates is complete by about the age of 3; (2) the process is latest in word initial position, affricates are most often fricativized (/t͡s/ → [s], /p͡f/ → [f]); (3) this substitution is more frequent for /p͡f/ than for /t͡s/; (4) at the same time fortition (/f/ → [p͡f], /s/ → [t͡s]) can also be observed and (5) the acquisition of affricates and of [ks] is achieved at about the same age, but governed by different strategies.

Finally, Schaner and Tonelli compare the acquisition of affricates in different languages, and suggest some explanation for their results. German and Czech children acquire affricates much before their English, Japanese, Maya or Italian pals. The speed of acquisition varies according to place of articulation as well, alveopalatal affricates (e.g., [ʧ]) are first, then come alveolar ones (e.g., [t͡s]), and labials (e.g., [p͡f]) are the hardest to acquire. Fricativization as an affricate substitution process is not characteristic of English, but does occur for Czech and Maya children, and it is especially common in every language word initially. The authors try to explain this by the smaller importance of the stop phase in this position, elsewhere it has an important role in perception: it abruptly cuts the signal, which is not the case in word initial position. Referring to Dogil–Jessen (in press), they claim that the more frequent fricativization of /p͡f/ is a result of the fact that the stop element of this affricate is less salient than that of /t͡s/. Hypercorrection should be the cause of the affricatization of word initial fricatives, which claim is supported by that it is done only by children who have already started acquiring affricates. The cluster [ks] was never found substituted either by fricativization or by a stop, which seems to prove its status of being a cluster of two phonemes as opposed to affricates.

The ten papers presented here cover a wide spectrum within linguistics leaving, nevertheless, fields about which Natural Theory has nothing to say. We have also seen that the possible claims the theory can make about a certain phenomenon are moving on a wide scale again, and that can be rather unfavourable of a theory. Although we may feel the limits, the axioms and theorems forced on us by a more constrained theory annoying, they are often more helpful in instigating new thoughts.

Péter Szigetvári

References

Brockhaus, W. 1990. Colourful leagues: A government phonology approach to final obstruent devoicing in German. In: UCL Working Papers in Linguistics 2: 270–97.

Dogil, G.–Jessen, M. (in press). Phonologie in der Nähe der Phonetik. Die Affrikaten im Polnischen und Deutschen. In: Prinzhorn, M. (ed.): Phonologie. Linguistische Berichte: Special Issue 2. Westdeutscher Verlag, Opladen.

Donegan, P.J.–Stampe, D. 1983. Rhythm and the holistic organization of language structure. In: CLS 19: 337–53. Parasession on the Interplay of Phonology, Morphology & Syntax.

Dressler, W.U. 1984. Explaining natural phonology. In: Phonology Yearbook 1: 29–52.

Gimson, A.C. 1989. An Introduction to the Pronunciation of English (4th edition). Edward Arnold, London.

Harris, J. 1994. English Sound Structure. Blackwell, Oxford.

Jones, D. 1956. The Pronunciation of English (4th edition). Cambridge University Press, Cambridge.

Keating, P. 1984. Phonetic and phonological representation of stop consonant voicing. In: Language 60: 286–319.

Lisker, L.–Abramson, A.S. 1964. A cross-language study of voicing in initial stops: acoustical measurements. In: Word 20: 384–422.

Schein, B.–Steriade, D. 1986. On geminates. In: Linguistic Inquiry 17: 691–744.

Temporal Factors in Speech: a Collection of Papers. Ed. by Mária Gósy. A Magyar Tudományos Akadémia Nyelvtudományi Intézete / Research Institute for Linguistics, Hungarian Academy of Sciences, Budapest 1991, 183 pp.

Temporal Factors in Speech (henceforward TFS) is a collection of seven—sometimes very different—papers, each of which approaches the temporal aspect or the "timing" of the phonetic/phonological representation. Two of the essays are primarily phonological (Tamás Szende: "On temporal speech distortions of present-day Hungarian" and Péter Siptár: "Fast-speech processes in Hungarian"), one is a study in historical phonetics/phonology (András Vértes O.: "The antecedents of quantitative versification in the historical phonetics of Hungarian"), three of them are phonetic studies—two of which (Olaszy's papers) directly or indirectly concern speech synthesis (Mária Gósy: "The perception of tempo", Gábor Olaszy: "The inherent time structure of speech sounds", and Gábor Olaszy: "Timing algorithms in the MULTIVOX automatic multilingual text-to-speech system", and one is a paper on language acquisition/child language (Mária Laczkó: "The interrelation of articulation rate and pauses in children's speech"). The papers in the volume are the result of research carried out in the Phonetics Laboratory of the Linguistics Institute of the Hungarian Academy of Sciences. All the papers focus on Hungarian.

I will proceed to discuss the papers—the space and attention devoted to individual essays reflect my bias and expertise, which are admittedly phonological.

Tamás Szende's paper has a dual objective: on the one hand he "make[s] an attempt at defining word-level phonological representation in an abstract form so as to include also a T component covering temporal pattern distortions" (1), on the other hand he discusses lenition (reduction, one of its subtypes, in detail) and its effects on the timing of the utterance. Accordingly, his paper divides into a theoretical and an analytical/descriptive part.

In the first part Szende argues that post-Sound Pattern of English phonological models cannot properly handle the timing aspect of phonological representations and proposes a "universal (word-level) P[honological] R[epresentation]" which consists of four strata or layers:

(1)
- i. ABSTRACT & GENERAL: "the lexical-morphosyntactic architecture of the PR" (6)
  szóval //#soː(-)#/ + /#-Cɒl#//
- ii. ABSTRACT & INDIVIDUAL: "the strictly phonological architecture of the PR" (6)
  szóval /soːvɒl/
- iii. CONCRETE & GENERAL: "the standard phonetic architecture of the PR" (6)
  szóval [soˑvɒl]
- iv. CONCRETE & INDIVIDUAL: "the actualised architecture of the PR" (6)
  szóval nem [[sa'nẹmː]]

(1iii) is called the "first grade", i.e. the "next-to-phonetic phonemic representation" (1) and is regarded as the input to lenition. There are some minor inconsistencies/inadequacies in Szende's illustration of his argument,[1] but the most interesting problem is the argument itself, viz. the critique of post-SPE phonology and the proposed solution: the multi-levelled representation.

Szende's argument against the way in which the temporal aspect appears in phonological theory represents a very interesting opinion even if one does not share his views. There is one point though where his treatment, I think, is a bit unfair to current phonological theory. He claims that both in traditional and post-SPE theories the phonological representation mirrors time in two ways: "(i) by a linear concatenation of phoneme size units [...] and (ii) by defining the actual position a phoneme size segment takes along the lines of the

[1] In the middle of page 1 (lenited) *szóval* is transcribed as [soˑvɔl], but under (4) on the next page the middle portion of the same word is represented as [oˑvɑ] at the same level. The difference in the quality of the second vowel is apparently accidental.

On page 3 Szende mentions a shortening rule which applies in some monosyllabic root morphemes and turns a long root vowel into a short one if the root is followed by a vowel-initial suffix provided the suffix vowel is [−high] and the root vowel and the suffix vowel are not identical. In Szende's formulation $(C)V{:}C \rightarrow (C)V_1CV_2$, $V_1 \neq V_2$, $V_2 = [-\text{high}]$. In actual fact this rule (which has heavy morphological conditioning) sometimes appears to apply even if $V_2 = /i/$ *t*[y]*zifa* 'firewood', *vas*[u]*ti* 'of the railway', [u]*tinform* ⟨proper name⟩.

The (complex) symbol (−) appears in representations at the "lexical-morphosyntactic level" (6). It is never explained in the text and has no standard interpretation.

Although it is explicitly pointed out that "in the given case the number of syllables is shown by the number of vowels" (4), this appears not to hold true of the example that the sentence cited refers to: \$laːʃ\$ʃ\$ɑ\$ (4). If dollar signs denote syllable boundaries, then this form consists of three syllables laːʃ, ʃ and ɑ.

'short/long' opposition whenever such a distinction is significant in the given language" (1). Then he goes on to point out that "restricting PR descriptions to the above-mentioned aspects of time results in inappropriate descriptions" (1). In my view point (ii) above is necessarily true of any phonological model (if only two, and not more than two, distinctive temporal distinctions exist universally). Point (i), however, is not true of autosegmental models in general, and the moraic theory (e.g. Hyman 1985; Hayes 1989) in particular. The reason why it is untrue of autosegmental phonology, is that the notion of the segment is drastically reinterpreted in these models (cf. van der Hulst 1984; Goldsmith 1990). As any feature may appear on a separate tier (and in most current views of feature geometry all of them do, cf. McCarthy 1988) and behave like a segment, all that remains of the traditional or SPE segment is the "timing slot" or X position on the timing tier (skeleton) to which the featural content and the suprasegmental structure (the syllable constituents, the grid, etc.) are associated. It is difficult to conceive the timing slots as "phoneme size units" because a single slot may be associated with two units on another tier and vice versa, a single timing slot may be associated with two units on some other tier. Moraic phonology goes further because it even dispenses with the timing tier. Root nodes (segmental features) are connected directly to the syllable node if they are in onset position or to morae which are dominated by the syllable node. In such a model length (time) is expressed by the morae, which are not segment size units.

The other problematic point in the theoretical part of the paper concerns the levels described above. In Szende's view there are four levels. The "deepest" is the morpho-phonological one (his "abstract and general") and the "shallowest" is a phonetic level (his "concrete and individual"). One wonders what the status of the two intermediate levels is. Are they phonological or phonetic? If we take "the first grade" to be phonetic—as is suggested by the description on page 6 according to which it is "the standard phonetic architecture",[2] then there is a phonological level (his "abstract and individual") wedged between the morpho-phonological level and the phonetic one(s). If this is so, then an argument against Halle's rejection of a phonemic level (the well-known duplication problem, cf. Halle 1959) would have been in order—all the more so because the first grade does seem (classical) phonemic (even if we keep the classical phoneme, it does not mean that we must/can keep the phonemic level—cf. Schane 1971).

The second part of Szende's paper deals with lenition in general, and analyses reduction, a subtype of lenition in detail. Apart from the Hungarian data, what is of importance here is what counts as lenition in general. This question is all the more pertinent because it allows a comparison with the next paper in the volume (Péter Siptár's article which also discusses lenition processes).

According to Szende, the main lenition types are the following (13–14): a) reduction, b) deletion (partial disappearance of segmental content), c) dropping and truncation (complete disappearance of segmental content), d) reduction over a sequence, e) sequence size truncation, and f) coalescence and fusion. In Siptár's paper the lenition/weakening processes discussed are i) deletion, ii) assimilation, and iii) reduction (this last type is a bit of a mixed bag because it includes a "miscellaneous" subtype). As can be seen, lenition means a wide variety of "things" and varies from author to author. Sometimes different types are very difficult if not impossible to distinguish. In Siptár's paper, for instance, reduction means (among other things) vowel shortening and degemination. The question is in what way is vowel shortening different phonologically from vowel deletion, or degemination from consonant deletion. One could say that these processes are simply deletions which take place in the environment of an identical segment. What is the difference between the subtypes

[2] Although this seems to be in contradiction with the description on page 1 according to which it is the "next-to-phonetic *phonemic* representation" (the italics are mine).

then? Not only are the subtypes of lenition uncertain, but it is hard to identify lenition itself. Szende says (10–11): "Lenition [...] does not simply mean the mirror processes of fortition: epenthesis[3] for example can also function as simplifying pronunciation, and thus a lenitive process". If one and the same process can be fortition and lenition, how do we know when it is a fortitive process and when it is a lenitive one? What is the theoretical status of lenition then? Neither author discusses this problem. Lenition is certainly a useful concept, but in order for it to be more than a handy metaphor, it has to be defined and identifiable in a formal way. The problem is certainly not a trivial one, but I am convinced that Harris (1990) is absolutely right in pointing out that it is not possible to provide a complete pretheoretical definition of lenition. Thus, defining lenition by simply listing the processes does not seem to be a tenable approach to the problem.

Péter Siptár's paper contains a thought-provoking theoretical discussion of what counts as fast speech and an excellent overview of fast speech processes in Hungarian. In what follows I will take up one theoretical problem and comment on a few of the specific processes Siptár analyses.

Siptár discusses Kaisse's idea that there are two blocks of rules in the postlexical component (Kaisse 1985): external sandhi rules, which have syntactic/morphological/lexical conditioning and fast speech rules, which are rate dependent and strictly phonological. In Kaisse's model the external sandhi block precedes the fast speech one. Siptár points out that Voicing Assimilation in Hungarian seems to be a problem, because it is postlexical, it has no lexical/syntactic/morphological conditioning (= it cannot be an external sandhi rule), but it is unaffected by the rate of speech (it cannot be a fast speech rule). As there is no third block the status of the rule is obscure. He suggests that there might be a way out: we can say that the rule is an external sandhi rule which is conditioned by "an empty set of syntactic etc. conditions" (35).[4] He also points out that there is a way in which this analysis can be disproved: all we have to do is find a rule that we know is a fast speech rule and which feeds (and therefore precedes) Voicing Assimilation. He then goes on to show that Cluster Simplification (e.g. *fogd meg* [fogmɛg]):

(2) $\begin{bmatrix} \text{+cor} \\ \text{- strid} \end{bmatrix} \rightarrow \emptyset \; / \; [\text{+cons}] \_\_ \; (\#) \; [\text{+cons}]$

is not such a rule because—although it is a fast speech rule—as forms like *oszd meg* [ozmɛg] show, Voicing Assimilation must precede Cluster Simplification (the reverse order would give *[osmɛg]). Siptár draws the conclusion that "Ellen Kaisse's conception of the internal organization of postlexical phonology is confirmed—or at least not refuted—by the Hungarian rules considered here" (36). Apart from the fact that there is a world of a difference between **confirming** and **not refuting**, Siptár does mention a rule later which actually disproves the above-mentioned organisation of the post-lexical component. The fast speech deletion of (high) vowels in doubly open syllables (38) (e.g. *öcsike* [øtʃkɛ]):

(3) $V \rightarrow \emptyset \; / \; VC \_\_ CV$

appears to be a rule which feeds Voicing Assimilation: *zaciba* 'to the pawnshop' [zɔdzbɔ], *radikális* [rɔtkaːliʃ] 'radical'. This rule is a fast speech rule which must precede Voicing As-

[3] Which neither author lists as a lenition process incidentally.

[4] I do not like this proposal (how do you distinguish an empty set of syntactic, etc. conditions from the lack of a set of syntactic, etc. conditions?), but this is not the point I want to raise here.

similation (the reverse order would give *[zɔtsbɔ], *[rɔdkaːliʃ]). This order seems to disprove Kaisse's hypothesis.

Some of the fast speech processes Siptár analyses have interesting theoretical relevance. He discusses two deletion rules whose probability of application increases **if the segments flanking the deleting segment are identical.** One of them is rule (3) mentioned above (*add ide* 'give it to me' [ɔdidɛ] → [ɔdːɛ]), the other is Intervocalic Consonant Deletion (*lehet* 'may be' [lɛhɛt] → [lɛɛt]):

$$(4)\quad \begin{bmatrix} \text{C} \\ \text{+voice} \\ \text{– strid} \end{bmatrix} \rightarrow \emptyset \;/\; \text{V} \_\_ \text{V}$$

These processes can be used in an argument against **antigemination** proposed in McCarthy (1986) who uses the Obligatory Contour Principle (a putatively universal constraint which bans identical adjacent elements on a melodic tier) to explain why deletion is often/usually blocked in natural languages if it would create a sequence of identical elements. Both the above rules have anti-antigemination effects, which disconfirms McCarthy (1986) and supports Odden (1988) who argues against antigemination. There really is no easy way out: both rules produce (actually favour the production of) sequences of identical elements on the same tier—there is no evidence that vowels and consonants were on different tiers at any stage of the derivation (let alone postlexically) in Hungarian. One possible "rescue operation" for antigemination would be to say that the OCP does not hold postlexically in Hungarian.

Another interesting theoretical issue which Siptár's analyses bear on is that of resyllabification. Siptár discusses a rule which deletes the liquids /l, r, j/ in coda position: *fel kell jönnöd* 'you must come up' [fɛːkɛːjønːød].[5] /l, r, j/ are "saved" by resyllabification in examples like *fel akarok jönni* 'I want to come up' [fɛlɔkɔrokjønːi] (and not *[fɛːɔkɔrokjønːi]—they are no longer in the coda if another vowel follows, but are resyllabified into onset position and thus are not affected by the rule. Interestingly enough, resyllabification happens **across a word boundary** although there are other processes in Hungarian which are sensitive to syllable structure but in which resyllabification is not permitted across a word boundary. /h/-deletion and epenthesis are rules of the latter kind. In /h/-deletion an underlying /h/ remains phonetically unrealised in coda position (e.g. *cseh* 'Czech' [tʃɛ]) and is not "saved" by resyllabification into onset position if the next word begins with a vowel (*Cseh akarok lenni* 'I want to be Czech': [tʃɛɔkɔroklɛnːi] and not *[tʃɛhɔkɔroklɛnːi]). Epenthesis breaks up unsyllabifiable final clusters as in *bokor* 'bush' /bokr/ → [bokor], but does not come into play if the stem-final consonant becomes syllabifiable by becoming an onset because a vowel-initial suffix follows *bokr-on* 'on a bush' [bokron]. However, the same stem-final consonants are not syllabified into the onset of a following vowel-initial word: *a bokor és a virág* 'the bush and the flower': [ɔbokoreːʃɔviraːg] and not *[ɔbokreːʃɔviraːg]. Thus, in some cases resyllabification appears to be possible and in other cases it does not. However, this contradiction is only apparent because the difference in behaviour can be explained. It is not due to the identity of the segments that are potential targets of resyllabification, but to the nature of the different rules. Epenthesis and /h/-deletion are **lexical** rules whereas Liquid deletion is **postlexical.** Thus, we can permit resyllabification postlexically because the segments targeted by /h/-deletion and Epenthesis have been deleted or have become syllabified in the lexical phonology and hence will not be affected by the postlexical application of resyllabification.

[5] [ɛː] is the result of compensatory lengthening.

Siptár's data are (according to this native speaker's intuition) surprisingly sound. I have found only one rule that he discusses which in my view is far more complex than he makes it out to be. The Full Assimilation of intervocalic clusters in both his formulations[6] completely assimilates the first member of an intervocalic cluster to the second in fast speech: *bizottság* 'committee' [bizoʃːaːg], *kiiktat* 'eliminate' [kiitːɔt]. While these examples are correct the rule seems to overgenerate: it fully assimilates all intervocalic clusters while some clusters do not seem to behave in this way:[7] compare *iktat* 'file⟨verb⟩' [itːɔt] and *kajtat* 'ransack' *[kɔtːɔt]. I do not pretend to know how the rule really works, but it seems that the sonority of the segments has a role: the greater the sonority difference between the two segments, the less likely it is that the rule will work: *hakni* 'gig' *[hɔnːi], *Mátra* ⟨proper noun⟩ *[maːrːɔ].

Mária Gósy's well-written and informative paper on the perception of tempo begins with a thorough review of the relevant literature. Then she reports on a series of experiments whose aim is the "analysis of (i) perceptual characteristics of speech tempo, (ii) the connection between tempo perception and speech comprehension [...and] (iii) the interrelatedness between the speaker/listener's own tempo and his tempo perception and comprehension" (64). In the experiments Hungarian native speakers were used as test subjects.

In the first set of experiments native speakers were asked to make judgements about the tempo of different versions of an artificially generated sentence which only differed in terms of F0 modulations (all other parameters were identical). In the listening tests the native subjects were asked to compare the speed of the sentences heard to their general idea of how fast Hungarian speech was. Gósy found that "(1) [...] Hungarian speech has accelerated over the past few decades; (2) melodic structure influences tempo judgements [...]; (3) utterances containing a melodic peak or a rising pattern are perceived as faster than descending ones; (4) monotonous utterances are judged [...] to be slower than the others up to a certain tempo value, whereas above that value [...] they are judged faster than the corresponding non-monotonous utterances" (76).

In the second set of experiments test subjects had to decide if the sentences they heard were true or false. The sentences they had to make judgements about were sometimes true, sometimes false and their tempo was manipulated by Varyspeech. There were four sets of sentences which only differed in tempo (normal, quick, very quick, very slow). The working hypothesis was that understanding should become more and more difficult as articulation time is reduced. This was disconfirmed by the experiment. Data showed significant differences only when very slow and very quick sentences were compared, otherwise no such differences were found.

Gósy's last set of experiments targeted the relation between tempo production and tempo perception. She points out that the usual view is that there is "a very close connection between the speaker's own rate of speaking and his perception and comprehension with respect to speech tempo" (82). The experiments conducted showed that "there is a slight but definitive difference of tempo perception among the subjects with various speech tempo production" (88) but "[t]he extreme speakers, namely the 'slow' and 'fast' speakers tend to behave perceptually similarly while 'moderate' and 'rapid' speakers tend to differ from the previous two groups".

Gábor Olaszy's first paper (The inherent time structure of speech sounds) is an acoustic approach to the inherent time structure (ITS) of speech sounds. The ITS is revealed in

[6] An autosegmental formulation (in which the root note of the second segment spreads backward to the X-slot of the first segment while the root node of the first segment delinks) would be theoretically more desirable than Siptár's formulations, but this is not the point here.

[7] Siptár does show that postconsonantal /h/ behaves in a different way.

the variations both in the frequency and energy components of the spectral structure of speech. Olaszy carried out an examination of the ITS of speech sounds using the analysis by synthesis method. The bulk of his paper is a detailed description of speech sounds in terms of time structure. Classification along these lines does not mirror the traditional division of speech sounds into vowels and consonants, but establishes another major cut. Speech sounds fall into two fundamental classes: 'simple structure' (SS) sounds (those that can be pronounced continuously: monophthongal vowels, spirants, most sonorants) and 'complex structure' (CS) sounds (those that cannot be pronounced continuously: affricates, stops, tremulants, diphthongs). As Olaszy points out tremulants are CS in spite of the fact that they can be prolonged.

After establishing this basic division he first examines the time structure of SS sounds in whose pronunciation "the frequency structure does not change; hence, in order to determine their time structure only the intensity structure must be examined" (110). He distinguishes an initial and a final phase and characterises each in detail. Then he goes on to examine CS sounds, which are characterised by varying intensity and frequency structure. Stops, affricates and tremulants are discussed in detail. The paper ends with several plates showing microintonation as a function of time in CV and VCV combinations and the time structure of CVC combinations in Hungarian.

Olaszy's second paper discusses the representation of timing in MULTIVOX, a multilingual system for text-to-speech synthesis. The paper describes the algorithms used for setting duration and other timing structures. First MULTIVOX itself is characterised, then the timing of segmental and suprasegmental units. The paper ends with a description of the timing options for the user.

Mária Laczkó's paper "investigate[s] children's speech production by comparing their use of articulation rate, pause and grammatical structures" (140). She examined two age groups (four and fourteen year old children). After carefully laying out the method and the material used she shows that the results partially confirmed her original assumptions that "articulation rate and the occurrence of hesitations in spontaneous speech correlated with subjects' age such that younger children speak slower and older children faster" (140). She did indeed find that the experimental data supported the hypothesis that younger children speak slower than older ones, but she also found that the younger the child was, the fewer and/or shorter hesitations occurred in his/her speech. On the basis of the various experiments carried out she suggests that "a higher level of first language acquisition is characterized in speech production, not independently of a more developed and diversified set of thinking and speech planning processes, by a higher proportion of hesitations" (151).

András Vértes O.'s paper closes the volume. It is about the emergence of the metrical possibility for quantitative versification (the short (light) syllable–long (heavy) syllable opposition) in Hungarian. A syllable is heavy if a) either it has a long vowel, or b) it is closed (regardless of the length of the nuclear vowel).[8] In Early Ancient Hungarian vowels were predominantly short and syllables were predominantly open. Therefore, there were very few metrically heavy syllables. Vértes O. shows how this state of affairs began to change in Ancient Hungarian and changed completely by the 16th century. He discusses in detail the processes (vowel lengthening in different environments and vowel deletion and gemination rules that created consonant clusters) that produced heavy syllables (weight by position: cf. Hayes 1989). Vértes O.'s paper is very interesting and well-argued. It is a curious thing (and

[8] Vértes O.'s wording is somewhat confusing. He says "[a light syllable] must be open, or if it is closed, a vowel initial syllable must follow" (76). Of course, if a putatively syllable-final single consonant is followed by a vowel (or a single putatively syllable-final stop is followed by a liquid), then it is no longer syllable-final, but syllabifies with the next syllable and leaves the first one open.

has little to do with Vértes O.'s article) that the potential distinction between heavy and light syllables—which is used in poetic meter in Hungarian—has no role in the Hungarian phonological system (in the sense that there are no rules sensitive to it[9]).

Miklós Törkenczy

References

Goldsmith, J.A. 1990. Autosegmental and Metrical Phonology. Basil Blackwell, Oxford.

Halle, M. 1959. The Sound Pattern of Russian. Mouton, The Hague.

Harris, J. 1990. Segmental complexity and phonological government. In: Phonology 7: 255–300.

Hayes, B. 1989. Compensatory lengthening in Moraic Phonology. In: Linguistic Inquiry 20: 253–306.

Hulst, H. van der. 1984. Syllable Structure and Stress in Dutch. Foris, Dordrecht.

Hyman, L. 1985. A Theory of Phonological Weight. Foris, Dordrecht.

Kaisse, E. 1985. Connected Speech: the Interaction of Syntax and Phonology. Academic Press, New York.

Levin, J. 1985. A Metrical Theory of Syllabicity. Ph.D. dissertation. MIT, Cambridge, MA.

McCarthy, J.J. 1986. OCP effects: gemination and antigemination. In: Linguistic Inquiry 17: 207–63.

McCarthy, J.J. 1988. Feature geometry and dependency: a review. In: Phonetica 43: 84–108.

Odden, D. 1988 Anti AntiGemination and the OCP. In: Linguistic Inquiry 19: 451–75.

Schane, S. 1985. The phoneme revisited. In: Language 47: 503–21.

Vago, R.M. 1989. Compensatory Lengthening in Hungarian: Implications for Mora Theory. Ms. City Univeristy of New York.

[9] There may be some more-or-less marginal phenomena sensitive to syllable weight distinctions (cf. Vago 1989; Levin 1985), but for instance, stress assignment is insensitive to them.

Levels and Structures (Approaches to Hungarian, Vol.5)

Edited by István Kenesei

The series *Approaches to Hungarian* was started in 1985 with the object of providing the international audience with the latest research on Hungarian. Volume 5 contains selected papers presented at an international conference held in Szeged, Hungary, 1994. The studies on phonology, morphology, syntax and semantics represent a wide and complex range of topics including optimality theory in phonology, focus and wh-movement, semantics and/or syntax of the definiteness effect, aspect and preverbal prefixes.
The contributors to the volume are: Farrel Ackerman (UC San Diego), Michael Brody (University College, London, and Institute of Linguistics, Budapest), Casper de Groot (IFOTT, Amsterdam), Julia Horvath (Tel-Aviv University), Katalin É. Kiss (Institute of Linguistics, Budapest), Catherine O. Ringen (University of Iowa), and Robert M. Vago (CUNY, New York).
Although the actual discussions could not be reproduced fully, the conflicting views and distinct approaches are well represented in the articles. In less controversial fields, for instance phonology of morphology, a wider coverage is provided.

About the editor: István Kenesei, Professor of Linguistics at the Department of English at József Attila University, Szeged; editor of all previous volumes of Approaches to Hungarian

Readership: linguists interested in theoretical studies or various aspects of Hungarian

In English, 1995, approx. 330 pages, 24x16 cm,
Paperback, USD 19.00
ISBN 963 482 051 4

Published by JATE
Distributor: Akadémiai Kiadó, Budapest

Akadémiai Kiadó, Budapest

Etymologisches Wörterbuch des Ungarischen
6. Lieferung

Editor-in-chief: Loránd Benkő

Das "Etymologische Wörterbuch des Ungarischen", dessen 6., abschließende Lieferung 1995 erscheinen wird, ist die organische Weiterentwicklung des in den Jahren 1967–1984 erschienenen vierbändigen "A magyar nyelv történeti-etimológiai szótára". In vielen Fällen wurden dessen Ansichten korrigiert, neue Etymologien vorgeschlagen, die Erstbelege vieler Wörter mit Hilfe neuer Quellen vorverlegt, neue Wortvarianten mitgeteilt und neue Stichwörter aufgenommen. Das Werk enthält bei ca. 10 000 Stichwörtern annähernd 50 000 ungarische wortetymologische Angaben, die über die hungarologischen Bezüge der Herkunftsproblematik des ungarischen Wortschatzes hinaus auch die Gesichtspunkte der Finnougristik, Turkologie, Indogermanistik, Romanistik und weiterer Bereiche erfassen. Es ist nicht nur ein wichtiges Handbuch für Linguisten, Historiker, Literaturwissenschaftler, Ethnographen, Kulturgeschichtler, Pädagogen und Bibliothekare, sondern auch ein unverzichtbarer Beitrag zur internationalen etymologischen Literatur.

About the editor: Loránd Benkő, Professor and Head of Department at the Eötvös Loránd University, Budapest, member of the Hungarian Academy of Sciences

Readership: linguists, historians, researchers in history of culture, undergraduate students

In German, 1995, approx. 300 pages, 17x25 cm,
Paperback, approx. USD 39.00
ISBN 963 05 6846 2

Akadémiai Kiadó, Budapest

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 with the following specifications:

operation system: DOS 2.0 or later;
floppy disk size: 5.25 or 3.5 inch;
floppy disk format: DS,DD (360/720 Kbyte) or DS,HD (1.2/1.44 Mbyte);
text editor: XYWrite, Word or Word Perfect;
text file: ASCII file without the formatting commands of the text editor;
special characters: choose some character combinations and use them consequently, for example: ə=\schwa{}; ɔ=\openo{}.

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, MTA Nyelvtudományi Intézet, H-1250 Budapest, P.O. Box 19.

Acta Linguistica Hungarica

AN INTERNATIONAL JOURNAL OF LINGUISTICS

Volume 42, Numbers 3–4, 1994

1828

Akadémiai Kiadó, Budapest

Acta Linguist. Hung. HU ISSN 1216–8076

ACTA LINGUISTICA HUNGARICA

AN INTERNATIONAL JOURNAL OF LINGUISTICS

Acta Linguistica Hungarica publishes papers on the subjects of Finno-Ugric, Slavonic, Germanic, Oriental and Romance linguistics as well as general linguistics in English, German and French.

Acta Linguistica Hungarica is published in yearly volumes of four issues by

AKADÉMIAI KIADÓ

Publishing House of the Hungarian Academy of Sciences

H-1117 Budapest, Prielle Kornélia u. 19–35

Manuscripts should be addressed to:

ACTA LINGUISTICA HUNGARICA
H-1014 Budapest, Színház u. 5–7

Correspondence with the editors and publishers should be sent to the same address.

Orders may be placed with

AKADÉMIAI KIADÓ
H-1519 Budapest, P.O. Box 245

Subscription price for Volume 42 (1994) in 4 issues US$ 84.00, including normal postage, airmail delivery US$ 20.00.

Acta Linguistica Hungarica is abstracted/indexed in Current Content–Arts and Humanities, Arts and Humanitites Citation Index and Bibliographie Linguistique/Linguistic Bibliography

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). The contribution proper should begin on page 2.

CONTENTS

PREFACE

This issue is a selection of abridged versions of papers from the Fifth Symposium on Logic and Language, which took place in August 1994 in Noszvaj, Hungary. The Symposia on Logic and Language are organized every second year by the Research Institute for Linguistics of the Hungarian Academy of Sciences, the Theoretical Linguistics Programme of Budapest University (ELTE) and the Department of Symbolic Logic and Methodology of Science of Budapest University (ELTE). They are wonderful occasions for logicians and linguists to discuss issues relevant to both linguistic theory and logic.

As witnessed by the present volume, the areas of common interest are numerous. The formal analysis of the logical properties of natural-language meanings as well as information flow itself helps linguists understand linguistic phenomena such as the behaviour of **more and less specific noun phrases** (Ábel and Maleczki), of **generic sentences** (Cavedon and Glasbey), of **quantificational expressions** (Farkas, Szabolcsi) and of **aspectual categories** (McClure). On the other hand, the types of meanings and ways of reasoning that linguists have come to posit for the explanation of linguistic phenomena constitute challenges to the logician. For example, the logical devices needed to explain **common-sense reasoning about verbal meanings** (Asher and Sablayrolles), the **dynamic treatment of presuppositions** (Beaver, Pólos), **felicity conditions** in general (Fox), **analogy comprehension** (Klink and Royakkers) and **metonymy** (Verschuur) involve intricate questions about logical calculi.

The above classification is probably very misleading. The now well-established collaboration of linguists and logicians is a challenging enterprise for both disciplines, but never in just one direction or the other. As a matter of fact, as the topics listed above show, many other areas, especially **cognitive science** bear a very similar relationship to both linguistics and logic as the lat-

1216–8076/94/$ 5.00 © 1994 *Akadémiai Kiadó, Budapest*

ter two to each other. The overlaps are so overwhelming that any attempt to classify the papers would run against the very spirit of the close co-operation I just mentioned. Therefore, we decided to arrange the papers in an essentially random order (namely, by the **a tergo** alphabetical order of their first authors).

László Kálmán
Guest Editor

Acta Linguistica Hungarica, Vol. 42 (3-4), pp. 149-170 (1994)

THE PROGRESSIVE: JAPANESE VS. ENGLISH*

WILLIAM McCLURE

In this paper, I compare the behavior of the English progressive with what appears to be its closest equivalent in Japanese. Naturally, I suppose, we find that the Japanese progressive is different from its English counterpart. Specifically, while the English progressive intuitively means something like an event is underway but is not yet complete, the Japanese construction can actually be progressive or perfective depending on the verb. Such behavior is obviously not predicted by existing theories which have all been developed to account for English (Dowty 1979; Parsons 1990; Landman 1992). I believe, however, that a parametrized analysis of the two aspectual systems is possible, and I believe that such an analysis reveals much about the natural language structure of events.

The basic problem which must be addressed is found in (1). In Japanese, all non-stative verbs occur in what is commonly known as the *te-iru* form. The gerund of the main predicate combines with a form of the verb *iru* 'be (of animate beings)'. With some verbs, the resulting expression is a true progressive, equivalent in all basic respects to the English progressive. With other predicates, however, the *te-iru* form is perfective, with no progressive connotations at all.

(1) The *te-iru* form (contrast)

(a) *Oyogu* 'swim'
Ima oyoide-iru
Now swimming-be-IMP

'(W) is swimming now'

(b) *Sinu* 'die'
Ima sinde-iru
Now, dying-be-IMP

'(W) is dead now / has died'
NOT
'(W) is dying now'

* I would like to thank Gennaro Chierchia and John Whitman for comments on earlier drafts of this paper. Mistakes are of course my own responsibility. This research was supported in part by the University of Durham.

1216-8076/94/$ 5.00 © 1994 *Akadémiai Kiadó, Budapest*

For a verb like *sinu* 'die', the *te-iru* form is in fact the most common way to say that en event has already taken place, and the only way to talk about an imminent death in Japanese is to use expressions such as 'is about to' or 'looks like'. Further, this difference in meaning is so strong, that it is often the most basic distinction made by traditional Japanese grammarians (cf. Kindaichi 1976; Koizumi *et al.* 1989).

Further examples of both kinds of predicates are given in (2), where all of the predicates in (2a) are progressive under *te-iru* while all of those in (2b) are perfective.

(2) More predicates (processes vs. discrete changes)

(a) Progressive	(b) Perfective
oyoide-iru 'is swimming'	*sinde-iru* 'has died, is dead'
aruite-iru 'is walking'	*tuite-iru* 'has arrived, is here'
hasite-iru 'is running'	*owatte-iru* 'has ended, is over'
benkyoo-site-iru 'is studying'	*kite-iru* 'has come, is here'
matte-iru 'is waiting'	*dete-iru* 'has left, is sticking out'
yonde-iru 'is waiting'	*kiete-iru* 'has gone out, is out'

If we look for an intuitive characterization of the kinds of predicates exemplified in (2a, b), we might say that the verbs in (2a) are all processes while those in (2b) are all discrete changes. Walking, for example, is as a series of steps which might go on forever and which is generally performed by some kind of conscious being who decides to go for a walk. In contrast, a discrete change such as *arrive* is a simple before and after. When a train arrives, first it is not here, and then it is. The two states define a discrete moment in time: the moment of change. Some verbs in a language, of course, are neither processes nor changes. In (3a) verbs such as *aru* 'have' or *iru* 'need' represent states, simple facts about the world. In general, such predicates are incompatible with the *te-iru* form in Japanese. In (3b), we see that the same is true of English, and stative predicates such as *have* or *need* are generally incompatible with the English progressive.

(3) Stative predicates

(a) Japanese	(b) English
**atte-aru* 'is having'	*W is having a book
**ite-iru* 'is needing'	*The car is needing more oil

These kinds of intuitions about types of events immediately bring to mind the notion of inherent aspectual structure or Aktionsarten. In English, inherent aspectual structure plays a prominent role in the distribution of various kinds of temporal adverbs and in entailments of the progressive. Dowty (1979) gives a particularly complete analysis of these phenomena. As we see in (4a, b), process-like verbs are compatible with expressions such as 'for three hours' while discrete changes are compatible with expressions like 'take three hours to'. In addition, we get a clear contrast with the entailments of the progressive. In (4c), the patterns of a stative predicate are also given. Statives, like processes, are open-ended without well-defined final moments, and intuitively, they may be true for moments or open intervals of time. As such they are compatible with 'for three hours' and not with 'take three hours to'. As statives do not involve any kind of change, however, they are incompatible with the progressive, as we have already seen.

(4) English aspectual types

(a) Process

W studied for three hours but never did figure out quadratic equations

*It took three hours for W to study

(NOT: 'start to study' or 'finish studying something'

W is studying → W has studied

(b) Change

*The fire went out for three hours

It took three hours for the fire in the backyard to go completely out

The fire is going out → The fire has not yet gone out

(c) State

W had a cold for three hours (and then miraculously got better)

*It took three hours for W to have a cold

(NOT: 'to get a cold')

*W is having a cold

(cf. 'W has a cold')

In (5), we see the same sorts of test applied to Japanese. Moriyama (1988) has explicitly taken much of Dowty's work and applied it directly to Japanese. In addition, he has added a number of compounding tests. Also, notice that while the content of *te-iru* based distinction is different from the English progressive test, the resulting verb classes are identical.

(5) Japanese aspectual classes (Moriyama 1988)

(a) *Doosa doosi* ('activity verbs')
benkyoo-si-hazimeru 'begin to study'
benkyoo-tuzukeru 'continue to study'
sanzikan benkyoo-suru 'study for three hours'
**sanzikan kakatte benkyoo-suru* 'take three hours to study'
benkyoo-site-iru 'is studying' is progressive

(b) *Tassei doosi* ('achievement verbs')
kie-hazimeru 'begin to go out'
**kie-tuzukeru* 'continue to go out'
**sanzikan kieru* 'go out for three hours'
sanzikan kakatte kieru 'take three hours to go out'
kiete-iru 'has gone out' is perfective

(c) *Zyootai doosi* ('state verbs')
**ari-hazimeru* 'begin to have'
ari-tuzukeru 'continue to have'
sanzikan aru 'have for three hours'
**sanzikan kakatte aru* 'take three hours to have'
**atte-iru* is bad

In traditional terms, process predicates are known as activities, while discrete changes are achievements. States, of course, are simply states. All of the predicates in (2a) are therefore activities while those in (2b) are achievements. The basic question to be answered in this paper, then, is why are Japanese activities progressive under *te-iru*, Japanese achievements perfective, and Japanese statives incompatible? We could of course simply define each aspectual class in terms of how it behaves under *te-iru*. This is a common approach taken within the Japanese literature (cf. Kindaichi 1976), and it parallels the literature on English where stative predicates are often defined as those which do not allow the progressive (Lakoff 1965; Dowty 1979; Parsons 1990). Equally, we could stop here and posit two homophonous *te-iru* constructions, one of which is progressive and one of which is perfective. There is in fact evidence for such a hypothesis in Korean and in some dialects of Japanese where the progressive and perfective meanings are actually represented by different morphemes. Both of these approaches are unsatisfying, however, because they ignore the many parallels which do exist between the two languages. Many of the adverbial and compounding tests obviously transfer from one language to the other, and two out of three aspectual classes do in fact behave similarly under the two constructions. Such an approach also begs the question, why should just activities be progressive and achievements not? And why is the *te-iru* form of an achievement equivalent to the perfective and not some other aspect?

The approach taken in this paper rests on two assumptions: (i) whatever inherent aspectual structure actually is, it is represented in more or less identical fashion in all languages and (ii) a single morphological construction should have a unified semantics. The first assumption means that inherent aspectual structure is a semantic universal. While this is obviously an empirical question, such an assumption is broadly motivated by the principles of cognitive science. Things which happen in the real world are observed and their temporal properties noted. The results of these observations are then reflected in language. However this happens and regardless of the actual nature of the final representations, presumably all humans do it exactly in the same way. More specifically, this assumption is motivated by the clear parallels which obviously exist between English and Japanese as illustrated in (4) and (5). Similar kinds of aspectual classes obviously exist in both languages, and it should also be noted that the notion of aspectual class has long played a role in the discussion of Japanese verbs independent of work in English. Kindaichi (1976) gives a thorough description of Japanese phenomena along with an extensive list of references. As far as I know, Moriyama is simply the first person to integrate the two traditions within the Japanese literature. The second assumption, that a single morphological form should have a single semantic interpretation, is broadly motivated by a principle of economy. Such a principle presumably makes a language easier to learn and to remember, and in standard morphological investigations one begins with the assumption that bits of language which are identical in form are also identical in meaning. Further, the fact that the English progressive and the Japanese *te-iru* form actually do behave in a parallel fashion with respect to two out of three aspectual verb classes makes such an analysis even more desirable.

My overall strategy, therefore, is to provide a general semantics for each of the inherent aspectual classes. This follows from my assumption that such classes are semantically universal. I then give a semantics of both the English progressive and the Japanese *te-iru* form. As we would expect, it is the language specific constructions which are semantically different, but the difference is quite straightforward and might be characterized as a parameter. The contrastive behavior of English and Japanese then follows from the interaction of the language specific construction and the general semantics for each of the inherent aspectual classes. In the final section of the paper, I apply this analysis to the futurate progressive in English and to various perfective interpretations of the *te-iru* form in Japanese.

Having argued for the general nature of inherent aspectual classes, I now give a formal semantic definition of each eventuality type. To begin, we review Dowty's (1979) aspectual calculus. Dowty's system is designed to account for the syntactic and aspectual patterns discussed above in (4) and (5). Dowty (as well as most other people working on English) actually assume four aspectual classes, examples of which are given in (6).

(6) Aspectual types (Dowty 1979 and many others)

(a)	States	(b)	Achievements
	stink		die
	belong		break (intrans)
	seem		sneeze
	love		recognize
(c)	Activities	(d)	Accomplishments
	swim		build a house
	push a cart		write a letter
	wait		read a book
	write papers		perform a symphony

Accomplishments, in particular, are a hybrid class with the characteristics of both activities and achievements. One can read a book for three hours and, equally, it can take three hours to read a book. In his formal representations, Dowty assumes the existence of simple stative forms and a finite set of three aspectual operators: BECOME, DO, and CAUSE. Semantically, Dowty's operators are defined as outlined in (7).

(7) Dowty's aspectual operators
$\mathrm{BECOME}(\alpha) = 1$ iff $\neg\, \alpha$ then β
$\mathrm{CAUSE}(\alpha, \beta) = 1$ iff $\neg\, \alpha \rightarrow \neg\, \beta$ (modal conditional)
$\mathrm{DO}(x, \alpha(x)) \rightarrow$ volitional (x) (material implication)

A state α BECOMES when it comes into existence. Although Dowty's definition is given in terms of interval semantics, its intuitive meaning is obvious, and BECOME is in some sense a pair of states, before and after a moment of change. I reformulate this definition in terms of situations, but I make no significant changes to the intuition behind the BECOME operator. Dowty's

CAUSE operator is somewhat more abstract, but at its most basic level, one event CAUSE's another when the absence of the first 'necessitates' the absence of the second. The full semantics for CAUSE as given by Dowty actually involves counterfactuals, and I do not address the nature of causation in this paper. Finally, in Dowty's theory, DO turns out to represent mainly the volitionality of the subject, and it means that the subject is a sentient being, theoretically in control of the course of events. Consequently, DO does not have an interval semantic or model theoretic definition. Unlike BECOME, it is not clear that DO can be thought of as a situation or a set of situations. As such it is left to play a minor role in Dowty's theory.

The four verb classes are then derived from underlying stative forms combined with the appropriate aspectual operators: BECOME for achievements, DO for activities, and both DO and BECOME in combination with the connective CAUSE for accomplishments. As we see in (8a), a state is a simple situation with no aspectual structure. No operators are incorporated into its representation. An achievement such as *die*, however, is represented as in (8b). 'Dead' is a state. 'BECOME dead' means 'to die'. *Swim* in (8c) is then an activity which entails a DOing by somebody of something. Note that in the representation of an activity the argument role appears twice. In Dowty's semantics, 'W swims' means something like 'W DOes W's swimming' or 'W controls W's swimming'. The subject of an activity is in some sense both the Agent of the process and the Experiencer of it. Accomplishments are then represented as in (8d), with the process of W building CAUSEing a house to come into existence.

(8) Dowty's aspectual calculus
  (a) State
  'W stinks' = ˆstink(W)

  (b) Achievement
  'W dies' = BECOME(ˆdead(W))

  (c) Activity
  'W swims' = DO(W, ˆswim(W))

  (d) Accomplishment
  'W builds a house' = DO(W, ˆbuild(W,x)
  CAUSE BECOME(ˆhouse(x))

The operators in these representations are then used to explain the distribution of the various time adverbials and the behavior of the English progressive. They cannot, however, explain why an achievement in English can have a progressive interpretation while an achievement in Japanese cannot.

I should also be obvious at this point that I have made no mention of accomplishment-type predicates in (4) and (5). In point of fact, I do not believe that accomplishments exist as a separate aspectual class. Rather, it seems that under the right circumstances (and Dowty makes the same observation), any activity can have the completed interpretation associated with the expression 'take three hours to'. What are traditionally called accomplishments are simply activities which are more likely to be aspectually ambiguous. Further, this ambiguity is almost always associated with some kind of syntactic change—the addition of a direct object, the addition of a goal, or a shift from a nominal to a sentential subject. While accomplishments do have a number of unique syntactic features which distinguish them from activities and achievements, they do not seem to have any unique temporal features, and as such, do not need individual treatment under a semantics of aspect.

I turn now to my analysis. The semantics of event structures which I define is designed to incorporate most of Dowty's results wholesale. It also exploits the compositional event-based definition of Dowty's aspectual semantics, although the representations which I define should be considered a development of Parsons's (1990) event approach to aspect. Like Parsons, I assume in my semantics that verbs are predicates of eventualities with arguments fed in locally under government by means of θ-roles. This means that all verbs have the general eventuality structure outlined in (9), while (10a) is the specific representation for 'W lives in Durham'. The stative predicate *live in Durham* is predicated of the eventuality s for state and W is given the relevant θ-role. In this instance, s stands for a possible situation. 'W lives in Durham' is then defined as the set of all possible situations s such that s is a living in Durham situation which also contains W. W is then mapped into the syntax by whatever θ-role is associated with the predicate.

(9) General eventuality structure (Parsons)
Verb = V(e)
NP + Verb = V(e) & θ(NP,e)

(10) Examples

(a) Stative

'W is in Durham'

$= \{s \mid \text{in Durham}(s) \ \& \ q(W,s)\} = \lambda s\,[\text{in Durham}(s) \ \& \ \theta(W,s)]$

(b) Eventive (i.e. non-stative)

'W swims'

$= \{e : \text{swim}(e) \ \& \ \theta(W,e)\} = \lambda e\,[\text{swim}(e) \ \& \ \theta(W,e)]$

In Parsons's semantics, there is an implicit basic split between states and events (i.e. between stative and non-stative predicates) as shown in (10a, b). Unlike events, which can change or not, states obtain throughout some period of time. By definition, they are homogeneous and unchanging during a particular interval, and they do not occur in the progressive because Parson's HOLD operator is inherent to them. All non-stative eventualities are then represented by a general event argument e which may or may not HOLD because it may or may not be homogeneous or unchanging. As such, non-stative predicates can occur in the progressive.

While these differences between statives and non-statives are clearly valid, rather than stipulating such a basic distinction, it is surely more satisfying to have the distinction fall out of the theory. In the semantics which I define here, I assume, like Dowty, that states are basic. Non-stative predicates are then defined in terms of structured sets of states. In particular, changes are represented by pairs of states while processes are represented by sets of pairs of states. Processes are sets of changes. The structured sets which I define replace the generic state or event argument proposed by Parsons, so that the relationship between stative and non-stative predicates is directly represented in the aspectual semantics of each predicate. As there are three basic aspectual structures, there are three sets of states to be defined, where each increasingly complex set is composed of simpler sets. Other parts of the theory of aspect, and in particular the progressive, can then be reformulated in terms of these sets of states. No major changes are necessary, however, in the event semantics developed by Parsons, because all I am doing is replacing a simple event representation, with another, slightly more complex event representation. In particular, the basic hypothesis of Parsons's work that sentences of English contain subatomic quantification over events is untouched by the changes I am proposing. Rather, as the increased complexity is internal to each kind of eventuality, the additional structure simply enables us to distinguish between types of eventualities, and it is this ability to distinguish types of eventualities which makes it possible to explain the behavior of the English progressive and the Japanese *te-iru* construction.

Like Dowty, then, I assume that stative predicates are the most basic aspectual structure, and in (11) I define Σ, the very large set which contains all possible situations.

(11) States (Σ)
Σ = the set of all possible situations where $s \in \Sigma$ is a particular situation

In (12a) through (d), I give formal definitions for the relatively standard properties of statives which I assume. While this is somewhat belabored, as the stative is the basic building block of my semantics, its characteristics are obviously crucial.

(12) Stative properties
(a) Temporal trace (τ)
Given $s \in \Sigma$, a particular situation, $\tau(s) = i$

(b) Less than or equal to for intervals ($\leq$)
Given two intervals, i and i', $i \leq i'$ iff
the least upper bound of $i \leq$ the least upper bound of i'

(c) Subinterval relationship ($\subseteq$)
Given the intervals i and i', $i' \subseteq i$ iff
$\forall i^*$ s.t. $i^* \leq i$, $i^* \leq i'$ & $\forall i^*$ s.t. $i \leq i^*$, $i' \leq i^*$

(d) Subinterval property of statives (density)
$\forall s$, any situation, $\forall i'$, $i' \subseteq \tau(s)$, $\exists s'$ s.t. $\tau(s') = i'$ & $s' = s$

In (12a), the temporal trace τ of a particular situation s is defined as the moment or open interval of time occupied by s. Intuitively, situations may be true for moments or open intervals of time. Imagine a ball bouncing on a table. If we photograph just the moment of contact, we could argue that the stative 'The ball is on the table' is true, but just for that moment. If the ball were resting on the table, the sentence 'The ball is on the table' is obviously true for some interval of time.

Relationships between intervals are then subject to the properties of time. I assume here that time is a dense partial order defined on the 'less than or equal relation' ($\leq$). Then, given that 'less than or equal to' is well-defined for moments of time (i.e. 11:58 $\leq$ 11:59), in (12b) we can say that one interval is then 'less than or equal to' another interval if the least upper bound of the first interval is 'less than or equal to' the least upper bound of the second (which is to say, the first interval ends before the second one does). In (12c), this is taken one step further. One interval is a subinterval of another if all the intervals above and below the first are also above and below the second. Basically, one interval is a subinterval of another if the first begins after the second and ends before it. Finally, in (12d) we can define the subinterval property of statives, also known as density. Basically, if we say that a situation holds for a particular interval, we also want the situation to hold at every subinterval of that interval. Returning to the ball, if I claim that a ball sits on the table for a full minute, it is the case that at any moment of time during that minute, we would expect to be able to look at the table and see the ball sitting on the table. This ability to look inside of an interval and find what is in some sense a smaller version of the predicate is known as density.

In (13), I then define what I call a maximal state. Unlike the definitions in (12), the definition in (13) is specific to my semantics. It is also crucial.

(13) Maximal situation
$\forall$s, s is a maximal situation iff
$\forall$i' s.t. τ(s) $\subseteq$ i', $\neg$ $\exists$ s' s.t. t(s') = i' & s' = s

A maximal situation is one which is not a subinterval of any other identical situation. It is in some sense a largest situation, and like a least upper bound or a greatest lower bound, it is unique. For the purposes of my semantics, then, what I refer to as a basic situation is also the largest dense interval where the situation obtains without interruption. What it means is that the inside of a particular situation is inaccessible to the aspectual semantics, and its effect is to force the semantics to look at the largest possible manifestation of a single situation and not some arbitrarily chosen part of it. For the purposes of my semantics, statives do not have parts. They are internally dense but nevertheless unanalyzable wholes, and as such they are all uniquely characterized.

Given the definition of density in (12d), and the definition of maximality in (13), we might now want to redefine statives as shown in (14).

(14) Statives (Σ) [cf. (11)]
Σ = set of all possible situations where
$s \in \Sigma$ is a particular (dense and maximally large) situation

Turning now to achievements, recall that an achievement is intuitively defined by two states, one before and one after the moment of change. If we pair states at random as in (15), we obviously get natural pairs of situations such as <is not here, is here> (meaning 'arrive') as well as many completely unnatural pairs such as <love you, be American> (which might be thought of as weird but nonetheless possible kind achievement). We also get pairs of states which do not seem to define a change in any normal sense of the idea. Thus pairs of identical states as well as pairs of states which are not in the right order with respect to time seem to violate our most basic understanding of what a change is.

(15) Pairs of states (= Possible changes)
<is not here, is here>
<love you, be American>
*<is blue, is blue>
*<is today, is yesterday>

The set of natural changes (otherwise known as achievements) might therefore be defined as the set of pairs of different states where the first precedes the second in time. Such a definition is given in (16).

(16) Changes (C)
$X = \Sigma \times \Sigma$ (the Cartesian product of Σ with itself) is the set of possible changes where $c = \langle s\ s' \rangle \in X$ is an **achievement** iff $s \neq s'$ & $\tau(s) < \tau(s')$

The set of all possible achievements is represented by C, the Cartesian product of the set of situations with itself. While this is obviously a huge set, a pair of states is a well-formed achievement if and only if it is composed of two different states and the first state precedes the second in time. This definition obviously captures Dowty's original intuition that a BECOMEing is defined as one state followed by a another.

Turning finally to activities, I have argued intuitively that a process is an open-ended set of steps or changes. Swimming is a series of strokes through the water and running is a series of steps. We can now see, however, that moving from one step to the next is a kind of change. As such, a process is really an open-ended set of changes. It is not, however, just an object containing a lot of random changes. Rather, a process is defined by an open-ended set of changes which are all of the same type and are all well-connected in time and space. Thus swimming is typically understood as a series of strokes, but equally it may be conceived of as a series of half-strokes or a series of quarter strokes. It may also be represented by a series of points in the water. The length of the interval defining the sequence doesn't matter, as long as a regular partition is possible. It would be strange to define swimming as a random series of single strokes, half strokes, and quarter strokes. Further, it is important that these steps, however they are defined, be well-ordered with respect to time. When swimming laps in a pool, the second lap cannot be preceded by the lap three but followed by the lap one. Thirdly, it must be the case that the various steps of a process are properly linked. In particular, each step of a process is defined as an achievement where achievements are pairs of states. Two achievements are properly linked if the final state of the first achievement is identical to the initial state of the second achievement (i.e. the changes must be linked head-to-tail). Thus it would be odd to define a stair climbing where one walked up all the odd stairs first and then went back and walked up the even stairs. All of the stages are of the same type. Even if the steps are ordered in time, it would be weird because the steps are not ordered in space.

These requirements on the structure of an activity are all contained in (17).

(17) Process (Π)
Π: $\Re \rightarrow X$, a function from the real numbers into the set of possible changes, defines the set of possible processes where
$p = \{c_1\ c_2\ c_3 \ldots\} = \{<s_1\ s_2>_1\ <s_2\ s_3>2 \ldots <s_{n-1}\ s_n>_m \ldots\} \in \Pi$
is an **activity** iff
[1] $\forall n\ \tau(s_{n-1}) < \tau(s_n)$ &
[2] $\forall n,\ c_{n-1} = <s_{n-1}\ s_{n-1}'>\ \&\ c_n = <s_n\ s_n'> \rightarrow s_{n-1}' = s_n$ &
[3] $\exists\theta$ s.t. $\forall s_n\ s_{n+1}$, s.t. $s_n\ s_{n+1} \in p$, $\theta(s_n) = \theta(s_{n+1})$

A process p, represented as a set of changes, is a member of Π, the set of all processes. Π is a function from the real numbers (or some other sufficiently large infinite set) into the set C. It creates open-ended sets of changes which are well-formed activities if the individual changes are : [1] ordered in time and

[2] linked head-to-tail. In addition, the third clause of (17) requires that every state of a process have at least one θ-role in common. An activity must be manifested through a particular individual. This is not an aspectual requirement as such and serves only to link the aspectual semantics of an activity with its proper syntactic representation. It has no consequences for the semantics of the progressive which I define today.

The representation of *swim*, a well-defined activity, is given in (18). Note in particular how the various achievements of the activity are connected.

(18) Activity (*swim*)
$p = \{c_1\ c_2\ c_3 \ldots\} = \langle s_1\ s_2\rangle_1\ \langle s_2\ s_3\rangle_2\ \langle s_3\ s_4\rangle_3 \ldots$
c_1 = stroke; s_1 = at loc_1, s_2 = at loc_2
c_2 = stroke; s_2 = at loc_2, s_3 = at loc_3
etc.
Swimming is a linear process where the final position of each stroke is the initial position of the next stroke.

Activities, like states, are defined over open intervals of time. Further, while activities are not dense in the strict sense of a stative because they are composed of achievements, they are nonetheless characterized by a kind of subinterval property which parallels that of statives given above. Compare (19a) which is copied from (12d) above and (19b).

(19) Subinterval properties
(a) Of statives (density) [cf. (12d)]
$\forall$s, any situation,
$\forall$i', i' $\subseteq \tau$(s),
$\exists$ s' s.t. τ(s') = i' & s' = s

(b) Of activities
$\forall p = \{c_1\ c_2\ c_3 \ldots c_n \ldots\}$, a process
$\forall$i'm i' $\subseteq \tau$(p) and i' is pragmatically large enough,
$\exists$p' s.t. τ(p') = i' & p' = p

Thus, every subinterval of a stative is identical to the stative itself. This is a consequence of density. In contrast, only a subinterval of an activity which is pragmatically large enough will contain a predicate identical to the activity itself. Returning to swimming, if we compare a half stroke with a set of five or six strokes, the first is definitely not a swimming while the second probably is. The minimum length of the subinterval necessary to recognize the predicate is determined by the nature of the process in question. A process such as swimming laps obviously entails an interval which is at least one lap long and probably longer. Moving, however, contains tiny, instantaneous steps and can be true over a very small subinterval. It is also worth noting that what I have called the subinterval property of processes is found in one form or another in all investigations into aspect (cf. the Activity Postulate of Dowty 1979 or Landman 1992), and the difference represented by (19a, b) is in fact the only feature which actually distinguishes the aspectual structure of states and activities.

The basic aspectual structures defined so far are summarized in (20). These three structures are taken from the complete definitions found in (14), (16), and (17), respectively.

(20) Basic aspectual structures
stative = s, a particular situation
achievement = c = $<$s s'$>$
activity = p = $\{<s_1\ s_2>\ <s_2\ s_3> \ldots <s_{n-1}\ s_n> \ldots\}$

We can see that achievements are defined as pairs of states while activities are defined as open-ended sets of achievements. Recall that every stative in these structures is maximally large and is therefore unique. Further, single statives and activities are potentially unbounded while achievements are finite.

In the last part of this paper, I turn finally to a comparative analysis of the English progressive and the Japanese *te-iru* form. I argue that the semantic difference between these two constructions is minimal and could be reduced to the setting of a single parameter. I also adapt my analysis to account for both the futurate progressive in English and various perfective uses of the *te-iru* form in Japanese.

Having defined general semantic structures for statives, achievements, and activities, I now define the well-formed pieces of each eventuality type. I call these pieces segments. A segment ε is basically any continuous subset of the situations in a well-formed aspectual structure. The well-formed segments of all three aspectual types are defined in (21).

(21) Possible segments of each eventually type ($\varepsilon \subseteq P$)

(a) Stative s

$\varepsilon \in \{s\}$, the entire eventuality

(b) Change $<s\ s'>$

$\varepsilon \in \{s, s', <s\ s'>\}$

(c) Process $\{<s_1\ s_2> <s_3\ s_4> \ldots <s_{n-1}\ s_n> \ldots\}$

$\varepsilon =$ any continuous subset $\{<s_p\ \ldots s_q>\}$ of

$\{<s_1\ s_2> <s_3\ s_4> \ldots <s_{n-1}\ s_n \ldots>\}$ where $1 \leq p \leq q \leq n$

A particular segment may include all of the original situations in the original structure or it may include only one. Recall from (13) above that each situation of an event structure is defined to be maximally large. Single situations do not have internal parts which are semantically accessible and they cannot be interrupted. The segment represented by a single situation is therefore the smallest possible segment of a particular eventuality, and the only possible segment of a stative is the stative itself.

A final segment is then any segment which contains a final state. Intuitively, a final segment is all of the eventuality from a given point to the end. My truth conditions for the progressive rely crucially on the fact that the final segments for each kind of eventuality are defined by different sets of basic situations. In (22), we see that only achievements actually have unambiguously final segments.

(22) Final segments of each eventuality type (final $\varepsilon \subseteq P$)

(a) Statives s

final $\varepsilon \in \{s\}$, the entire eventuality

(b) Change $c = <s\ s'>$

final $\varepsilon \in \{s', <s\ s'>\}$

(c) Process $p = \{,s_1\ s_2> <s_3\ s_4> \ldots <s_{n-1}\ s_n> \ldots\}$

final $\varepsilon = \emptyset$

The single segment if the stative in (22a) is final, but it is also intitial. Activities are represented by open-ended sets of states. As such, they never have final states, and by extension, they never have final segments. In (21b), however, the segments s and $<s\ s'>$ are both final segments because they both contain a final state. In addition, the single situation s' is an unambiguously final segment.

With the definition in (21) and the notion of final segment given in (22), we now have enough structure to define the truth conditions for the progressive in both Japanese and English. I begin here with English. Intuitively, the English progressive is used when an eventuality is begun but is still incomplete. In terms of sets of situations, an eventuality is in progress during a given interval of time if some but not all of it its situations have been realized during that interval of reference. (23) gives my truth conditions for the progressive in English. The bracketed numbers used throughout this discussion are for reference only.

(23) The Progressive (English)
PROG(P) = 1 during the interval i iff
[1] $\exists\varepsilon$ s.t. $\varepsilon \subseteq P$ & $\tau(\varepsilon) \leq i$ &
[2] $\neg[\forall e^*$ s.t. $\varepsilon^* \subseteq P, \tau(\varepsilon^*) \leq \tau(\varepsilon)]$ &
[3] $\forall\varepsilon' [[\forall\varepsilon^*$ s.t. $e^* \subseteq P, \tau(\varepsilon^*) \leq \tau(\varepsilon')] \rightarrow \tau(\varepsilon') > i]$

One line at a time, the definition means the following. The progressive is true during the interval i if [1] there is a segment e of a predicate P which is manifested during or before i, the interval of evaluation; [2] it is not the case that this segment is later than all other segments of the predicate (i.e. ε is not a final segment); and [3] any segment which is a final segment is manifested after the interval of evaluation. The progressive is true for a particular interval of time if during that interval the eventuality has begun but is not yet complete.

How does the definition in (23) interact with the eventuality structures for each of the aspectual classes to predict the distribution and behavior of the progressive? First, English statives are never felicitous with the progressive because they contain only one segment. While line [1] can be satisfied, lines [1, 2] together can never be satisfied because the single situation of a stative is the smallest possible segment and it is both initial and final. It is therefore impossible to find both a non-final segment and a final segment because only one segment is defined. (23) therefore makes it unnecessary to stipulate that states do not occur in the progressive. Activities do not have well-defined final states by definition so final segments are also impossible to identify. The universal in line [3] is always well defined, and the progressive of an activity is true in an interval whenever the temporal trace of a non-final segment exists in that interval. Given the subinterval property of activities in (19b) above, the minimum length of the manifested segment is pragmatically determined. It must, however, be long enough for us to recognize the activity. As such, it is (19b) which gives us the fact that the progressive of an activity entails the perfective. The manifested piece is perfective while the entire predicate is progressive, but (19b) guarantees that the manifested piece and the entire predicate are indistinguishable from each other. Finally, as achievements are defined by sets of only two states, possible segments are few. Lines [1, 2] of the definition in (23) are satisfied by the two states which define an achievement, but the progressive holds only if the temporal trace of the earlier state is manifested before or during the interval of evaluation, while the later state is manifested after. The progressive of an achievement is felicitous only when the interval of evaluation actually cuts the achievement in half.

The English progressive is therefore true when part but not all of an eventuality is manifested. As states do not have parts, the progressive is never felicitous. Achievements and activities, on the other hand, may occur in the progressive but only under specific circumstances. In contrast to this, I believe that the *te-iru* form is basically a perfective aspect, requiring the manifestation of all final segments. In this it is intuitively the opposite of the English progressive. Truth conditions for the Japanese *te-iru* construction are given in (24). Again, the bracketed numbers are for reference only.

(24) The Progressive (Japanese)
PROG(P) = 1 during the interval i iff
[1] $\exists \varepsilon$ s.t. $\varepsilon \subseteq P$ & $\tau(\varepsilon) \leq i$ &
[2] $\neg[\forall \varepsilon^*$ s.t. $\varepsilon^* \subseteq P, \tau(\varepsilon^*) \leq \tau(\varepsilon)]$ &
[3] $\forall \varepsilon' [[\forall \varepsilon^*$ s.t. $\varepsilon^* \subseteq P, \tau(\varepsilon^*) \leq \tau(\varepsilon')] \rightarrow \tau(\varepsilon') \leq i]$

The Japanese progressive is true during the interval i if [1] there is a segment ε of a predicate P which is manifested during or before i, the interval of evaluation; [2] it is not the case that this segment is later than all other segments of the predicate (i.e. ε is not a final segment); and [3] every segment which is a final segment is manifested during or before the interval of evaluation. The progressive is true for a particular interval of time if the eventuality has begun at the interval of evaluation and, if possible, is also complete. The only difference between (23) and (24) is the status of the final segments in line [3]. This is the parameter. In English, final segments cannot be realized during the relevant interval. In Japanese, all possible final segments must be realized during the interval.

Like its English equivalent, the definition in (24) interacts with the structures defined for each aspectual type in a straightforward fashion to derive the necessary distributions. Again, statives never occur in the *te-iru* form because two unique segments are never defined. Lines [1, 2] together are never satisfied. Activities, then, do not have well-defined final segments by definition so they can never be complete. The universal in line [3] is always well-defined, and the *te-iru* form of an activity is felicitous during an interval whenever the temporal trace of a pragmatically long enough non-final segment exists in that interval. Significantly, the *te-iru* form of a Japanese activity is felicitous under exactly the same conditions as its English progressive equivalent. It is therefore not a surprise that an English activity in the progressive and a Japanese activity in the *te-iru* form refer to identical states of affairs in the world. Finally, as achievements are defined by sets of only two states, the set of possible segments is very small. Lines [1, 2] of (24) are satisfied by the two states which define an achievement, and *te-iru* holds if the temporal trace of both states is manifested during the interval i. The *te-iru* form of an achievement is possible only when all final segments have been realized. As such, it follows directly that the *te-iru* form of an achievement is perfective. The analyses in (23) and (24) therefore account for the basic puzzle presented in (1) at the beginning of this paper as well as the relationship between the progressive and each of the aspectual classes included in (4) and (5). While the English progressive and the Japanese *te-iru* form may look very different, they are in fact quite similar.

Finally, and briefly, I turn to the futurate progressive. In English, the progressive of a predicate can be used with a futurate interpretation as in (25).

(25) Futurate progressive
W is coming/going to come/will come tomorrow afternoon.
The sun ??is rising/is going to rise/will rise tomorrow morning.

There appear to be a number of semantic restrictions on the distribution of the futurate progressive (i.e. the action is generally intentional which is why 'the sun is rising tomorrow morning' is funny), but regardless of the exact nature of these restrictions, the Japanese *te-iru* form never has a futurate interpretation. This contrast follows directly from the definitions in (23) and (24). While the first two lines of (23) and (24) guarantee that an initial segment occurs sometime in the past, the manifestation of the complete event in English is restricted by the '>' symbol in line [3] of (23). As the end of the eventuality is always in the future, it is simply a matter of making the initial existential into some kind of intensional operator, putting the entire event into the future. One common way of doing this is to allow 'planning' stages at the beginning of an eventuality. Thus, planning to do something qualifies as having started the event. Such an approach explains the volitional restriction on the subject of a futurate progressive because an inanimate object cannot plan to do anything. In Japanese, however, the eventuality in question is restricted by the '$\leq$' symbol in line [3] of (24). It makes no sense to turn the initial existential into an intensional operator because *te-iru* is felicitous only when the outcome of the event is already manifested. Line [3] of the truth conditions for the *te-iru* construction guarantees that it can never be put into the future.

In contrast, however, note that the definition in (24) may be interpreted in such a way that the entire thing is perfective. This is represented in (26).

(26) Perfective *te-iru* [cf. (24)]
$\mathrm{PROG}(P) = 1$ during the interval i iff
[1] $\exists\varepsilon$ s.t. $\varepsilon \subseteq P$ & $\tau(\varepsilon) < i$ &
[2] $\neg[\forall\varepsilon^*$ s.t. $\varepsilon^* \subseteq P$, $\tau(\varepsilon^*) \leq \tau(\varepsilon)]$ &
[3] $\forall\varepsilon'$ $[[\forall\varepsilon^*$ s.t. $\varepsilon^* \subseteq P$, $\tau(\varepsilon^*) \leq \tau(\varepsilon')] \rightarrow \tau(\varepsilon') < i]$

In (26), the '$\leq$' symbols in lines [1] and [3] are both interpreted strictly, placing the entire relationship into the past. We therefore predict that the *te-iru* form of any Japanese predicate, achievement or activity, should have a perfective interpretation. We have already seen that achievements in Japanese are perfective under *te-iru*. Recall now that I argued against accomplishments as a class of predicates because all English activities seem to have the potential for a bounded interpretation. In Japanese, accomplishments as a class of predicates have never been recognized because all activities actually have bounded interpretations under *te-iru*. This interpretation is generally subordinate to the progressive interpretation (especially with an intransitive activity such as *hasiru* 'run'), but it is always available. Even in (1a) repeated here as (27), 'Ima oyoide-iru' is ambiguous and could mean that one is now swimming or that one has just now swum as far as some place. The analysis given in (23) and (24) explains this contrast in a very straightforward fashion.

(27) *Oyogu* 'swim' [cf. (1a)]
Ima oyoide-iru
Now swimming-be-IMP
'(W) is swimming now'
OR
'(W) has swum (somewhere) now'

To conclude, I believe that the contrasting behavior of the English progressive and the Japanese *te-iru* form as defined in (23) and (24), respectively, points to a larger difference between the two languages. Given that the imperfective/perfective distinction is basically binary, it should not come as a complete surprise that the 'in-between' construction (i.e. the progressive) in a given language actually falls on one side or the other. When we compare English and Japanese, I believe we find that the English progressive is basically an imperfective aspect while the Japanese *te-iru* form is basically perfective. This is of course confirmed in English by the futurate progressive and in Japanese by the progressive/perfective ambiguity which characterizes the *te-iru* form of activities. Contrariwise, the English progressive is never perfective, and the *te-iru* construction can never refer to a completely future event. This difference may reflect a general difference found in all languages: the progressive construction is parametrized as a perfective or imperfective form which is felicitous with all non-stative predicates. Or, it may simply reflect two constructions in two languages which are similar but not quite identical in meaning.

In either case, it is also clear that a more refined representation of event structure is necessary to capture the various kinds of behavior which result. A simple stative/non-stative classification is not enough because achievements and activities are both non-stative predicates, and they do not behave as a unified class. In this paper, I have proposed one possible representation based on the notion that all eventualities are constructed from particular sets of states. Achievements are pairs of states while activities are sets of pairs of states. From a cognitive perspective, we should expect eventualities to have the same structure in any language. How these structures are used is then a language specific question. The analysis given here offers insight, therefore, into the basic differences between Japanese and English aspectual behavior, as well as into the overall structure of events in natural language.

References

Dowty, D.R. 1979. Word Meaning and Montague Grammar. Reidel Publishing Company, Dordrecht-Boston.

Kindaichi, H. 1976. Nihongo Doosi no Asupekto [Aspect in Japanese Verbs]. Mugi Syooboo, Tokyo.

Koizumi, T. *et al.* 1989. Nihongo Kihon Doosi Yoohoo Ziten [A Dictionary of Basic Japanese Verb Usage]. Taisyuukan Syoten, Tokyo.

Lakoff, G. 1965. On the nature of syntactic irregularity. Ph.D. dissertation. Indiana University. (Published by Rinehart & Winston as Irregularity in Syntax (1970).)

Landman, F. 1992. The progressive. In: Natural Language Semantics 1: 1–32.

McClure, W. 1994. Syntactic projections of the semantics of aspect. Ph.D. dissertation. Cornell University.

Moriyama, T. 1988. Nihongo Doosi Zyutugobun no Kenkyuu [Investigations of Japanese Verbal Predicates]. Meiji Syoin, Tokyo.

Parsons, T. 1990. Events in the Semantics of English: A Study in Subatomic Semantics. The MIT Press, Cambridge.

Address of the author: William McClure
Department of East Asian Studies
University of Durham
Durham DH1 3TH
United Kingdom
e-mail: w.t.mcclure@durham.ac.uk

Acta Linguistica Hungarica, Vol. 42 (3-4), pp. 171-187 (1994)

ALL QUANTIFIERS ARE NOT EQUAL: THE CASE OF FOCUS*

ANNA SZABOLCSI

The traditional ideal of formal elegance requires that the treatment of scope and other quantificational phenomena be designed to be entirely general. It turns out, however, that natural language quantifiers do not behave uniformly. One possible reaction is to maintain the general formulations and supplement them with filters. A more explanatory strategy may be to devise specialized mechanisms that immediately give correct results. The following heuristics was proposed in Szabolcsi (1994a, 658):

(1) What quantifiers participate in a given process is suggestive of exactly what that process consists in.

Some of the cases in which this strategy has proven useful are the treatments of anaphora (Kamp–Reyle 1993), algebraic constraints on wide scope (Szabolcsi–Zwarts 1993), weak and strong readings of donkey sentences (Kanazawa 1994), branching readings of subject/object pairs (Beghelli *et al.* 1993), quantifier scope interaction (Liu 1990; Ben-Shalom 1993; Beghelli 1993; Dobrovie-Sorin 1993; Stowell–Beghelli 1994), quantifiers in pair-list readings (Szabolcsi 1994a). The present paper is concerned with another case of empirical interest.

Hungarian is known to be a language that wears its Logical Form on its sleeve: it has a number of surface syntactic positions whose sole function is to disambiguate scope and focus. In this paper I observe that each such position accommodates only certain quantifiers. The relevant classes are upward monotonic, group denoting, distributive, and cardinality quantifiers. In the spirit of (1), the constraints can be traced back to the semantics each position is associated with.

* This research was partially supported by NSF grant No. SBR-9222501. I thank László Kálmán for comments.

The paper first outlines the distribution of quantifiers in the "scope positions", then singles out one such position for discussion: the Focus position. Detailed discussion of the general picture is to be found in Szabolcsi (1994b).

1. Quantifiers in Hungarian surface structure

A substantial body of work by Hunyadi, Kenesei, É. Kiss, Szabolcsi, and others since the early eighties has established that surface order and intonation disambiguate scope in Hungarian.[1, 2] For instance, the following sentences are unambiguous:

(2) (a) *Sok ember mindenkit felhívott.*
many men everyone-acc up-called
'Many men phoned everyone'
many men > everyone

(b) *Mindenkit sok ember felhívott.*
everyone-acc many men up-called
'idem'
everyone > many men

(c) *Hatnál több ember hívott fel mindenkit.*
six-than more men called up everyone-acc
'More than six men phoned everyone'
more than 6 men > everyone

[1] For simplicity, I assume that in (2c) the postverbal universal is unstressed.

[2] Independent (cumulative) quantification is also possible and requires a special word order, whose syntactic analysis is presently unclear to me:

(a) *Kevés ember (mindig/gyakran) keveset végez.*
few men always/often little-acc accomplishes
'Few men tend to accomplish little (between them)'

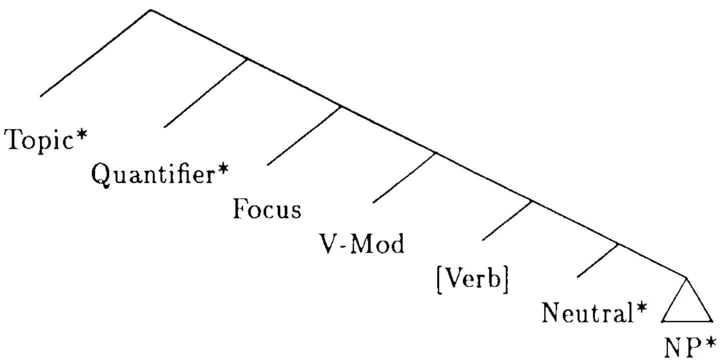
(b) *Kevés ember végez keveset.*
few men accomplishes little-acc
'Few men are such that they accomplish little'

(c) **Keveset végez kevés ember.*
little-acc accomplish few men

(d) *Mindenkit hatnál több ember hívott fel.*
everyone-acc six-than more men called up
'idem'
everyone > more than 6 men

Theoretically speaking, it is the occurrence in specific syntactic positions that defines the quantifier's scope. Simple syntactic tests distinguish the positions in (3); the postverbal position is assumed to be base generated. I use the traditional pretheoretical labels.[3] See É. Kiss (1991, 1994) for recent discussion, and Stowell–Beghelli (1994) for an interestingly similar proposal for Logical Form in English.

(3)



It is also well known that different noun phrases (DPs) occur in different positions, e.g., a Topic must be specific, and universals do not occur in Topic or Focus. However, no systematic investigation of these matters has been carried out to date. In what follows I examine a representative sample. Note that many DPs occur in more than one class; as we shall see, their interpretations vary accordingly.

[3] (i) Focus: strictly precedes the finite verb stem; if the verb has a prefix, prefix/verb inversion is triggered: *fel-hív* 'up-call' ⇒ *hív fel*; can be constituent-negated.
(ii) Verbal Modifier: in complementary distribution with Focus; cannot be constituent-negated. (Might be called non-contrastive Focus.)
(iii) Topic: sentence initial; can be followed by adverbials like *tegnap* 'yesterday'.
(iv) Quantifier: none of the above; can follow Topic and precede Focus/V-Mod.

I omit the position of Contrastive Topic (Left Dislocation), because it is in every respect like the Neutral position. Contrastive Topic is distinguished from Topic both by having a "scooped" intonation and by needing to be followed by Focus or negation.

(4) Type A: DPs that may occur in Topic and Focus positions: *Péter*, *Péter és Mária* 'P and M', *a fiú(k)* 'the boy(s)', *hat fiú* 'six boys'.

(5) Type B: DPs that may occur in Quantifier position: *minden fiú* 'every boy', *valamennyi fiú* 'each boy', *még Péter is* 'even Peter', *hat fiú is* 'as many as six boys', *Péter is* 'Peter, too', *Péter sem* 'Peter either', *semelyik fiú*[4] 'none of the boys', *sok fiú* 'many boys', *több, mint hat fiú* 'more than six boys (1)', *legalább hat fiú* 'at least six boys'.

(6) Type C: DPs that must occur in V-Mod unless Focus is filled; may also be pied piped to Focus if they have a contrastive constituent: *pontosan hat fiú* 'exactly six boys', *kevés fiú* 'few boys', *kevesebb, mint hat fiú / hatnál kevesebb fiú* 'less than six boys (1,2)', *legfeljebb hat fiú* 'at most six boys', *több, mint hat fiú / hatnál több fiú* 'more than six boys (1,2)', *KÉT fiú* 'TWO boys', *SOK fiú* 'MANY boys'.

(7) Leftovers: (i) DPs that may occur in Topic but not in Focus: *a legtöbb fiú* 'most (of the) boys' *valamely fiú* 'some boy;' (ii) DPs that may occur in V-Mod or Focus: *fiú(k)* 'boy(s), existential'.

The three main classes are semantically coherent. In Kamp–Reyle's (1993) sense, the noun phrases in Type A are set (group) denoters, and those in Types B and C, generalized quantifiers. However, Type B sides with Type A in English in supporting inverse scope and non-distributive *de re* readings (Farkas 1993; Stowell–Beghelli 1994), whence a three-way classification is needed:[5]

(8) (a) *Some man or other hit two boys.*
ok 'two boys > some man'

(b) *Some man or other hit every boy.*
ok 'every boy > some man'

(c) *Some man or other hit few boys.*
?? 'few boys > some man'

[4] Hungarian is a so-called negative concord language, and Szabolcsi (1981) argued that *se*-phrases are universals that must take immediately wider scope than negation. (Current literature on Italian makes comparable claims.) Similarly for *Péter sem*, the negative concord version of *Péter is*. The licensing negation may be preverbal or pre-Focus negation. Both their scope behavior and their ability to precede a filled Focus indicate that *se*-phrases, unlike real decreasing QPs, belong to Type B.

[5] The label "ok" is used to indicate that the given reading is possible, though not necessarily the only reading of the sentence.

(9) (a) *Some man or other imagined that two fools were smart.*
ok two fools *de re*; men do not vary

(b) *Some man or other imagined that every fool was smart.*
ok every fool *de re*; men do not vary

(c) *Some man or other imagined that few fools were smart.*
?? few fools *de re*

The tests motivating the three-way classification can be replicated in Hungarian. Szabolcsi (1994b) argues in detail for the following characterization:

(10) (a) Type A items are set (group denoters).

(b) Type B items are set (group) denoters that are prototypically associated with a distributive operator in their own clause.

(c) Type C items are cardinality quantifiers (not necessarily intersective).

Striking support for the proposed interpretations of Types B and C comes from DPs that occur in both Quantifier and Verbal Modifier positions, e.g., *több, mint hat fiú* 'more than six boys (1)'. In Quantifier position they offer up a set (group) for exclusively distributive predication; in V-Mod they count the atoms of the group / elements of the set denoted by the predicate:

(11) (a) *Több, mint öt fiú felemelte az asztalt.*
more than five boy up-lifted the table-acc
'More than five boys lifted the table (individually)'

(b) **Több, mint öt birtok körülöleli a kastélyt.*
more than five estate around-embraces the the castle-acc
'More than five estates surround the castle'

(12) (a) *Több, mint öt fiú emelte fel az asztalt.*
more than five boy lifted up the table-acc
'The number of boys who lifted up the table (individually) is five'
or
'It took more than five boys to lift the table (collectively)'

(b) **Több, mint öt birtok öleli körül a kastélyt.*
more than five estate embraces around the the castle-acc
'The group of estates that surrounds the castle has more than five atoms'

More generally, Szabolcsi (1994a,b) characterizes the mode of operation of Type A and Type B items as "predicate and ±distribute", and that of Type C items as "count". Note now that Types A and B contain only upward monotonic quantifiers.[6] In the spirit of (1), the reason is that only these can define a group (set) that serves as a target of predication. With this background, let us focus on focus.

2. Group denoters in free focus

2.1. Groups and exhaustivity

In accordance with current literature, I will distinguish free focus and modification by *csak* 'only'. When talking about the occurrence of DPs in the Focus position, I will by default have free focus in mind. *Csak* is discussed in section **3**.

Let us now recap the pertinent observations of section 1. I was claiming that only two types of DPs show up in the syntactic Focus position: those belonging to Type A, on their own right (13), and those belonging to Type C, when they have a constituent set into contrast (14c). Recall that the normal habitat of Type C DPs is the Verbal Modifier (VM) position, which differs from Focus in not supporting constituent negation.[7] Type B is excluded from Focus (15).

(13) (a) [$_F$ *Hat fiút]* *láttam.*
six boys-acc see-past-1sg
'It was six boys that I saw'

(b) *Nem [$_F$ hat fiút]* *láttam.*
not six boys-acc saw-I
'It was not six boys that I saw'

(14) (a) [$_{VM}$ *Hatnál kevesebb fiút]* *láttam.*
six-than fewer boy-acc saw-I
'I saw fewer than six boys'

[6] *Semelyik fiú* 'none of the boys' superficially contradicts the upward monotonic generalization. See note 4.

[7] As is indicated in (6), bare numeral indefinites like *hat fiú* 'six boys' belong to Type C iff their head noun is destressed. In this case they mean 'exactly six boys'.

(b) *Nem láttam hatnál kevesebb fiút.*
not saw-I six-than fewer boy-acc
'I didn't see fewer than six boys'

(c) *Nem [$_F$ HATNÁL kevesebb fiút] láttam, hanem*
not six-than fewer boys-acc saw-I but
[$_F$ ÖTNÉL kevesebbet].
five-than fewer-acc
'It was not fewer than SIX boys that I saw but fewer than FIVE'

(15) *[$_F$ *Minden fiút] láttam.*
every boy-acc saw-I
'It was every boy that I saw'

Parallel to these distributional facts the following interpretive contrasts can be observed:

(16) (a) Type A DPs receive an exhaustive interpretation in Focus. E.g. (13a) cannot be true if I saw anyone beyond six particular boys.

(b) Type C DPs do not receive an exhaustive interpretation in Focus: at best their numeral part is affected.

(c) There is no exhaustive interpretation outside Focus.

The question I wish to address in this section is why this is so. Recall the heuristics proposed in (1):

(1) What quantifiers participate in a given process is suggestive of exactly what that process consists in.

Since we know that Type A DPs are set (group) denoters, this tells us that exhaustivity in the requisite sense must be applicable only to sets (groups). Below I show that this is indeed the case.

Prior to proceeding to the details, recall that there are two essentially different approaches to free focus in the literature. The first assumes that free focus makes a semantic contribution (affects the presupposition and/or assertion of the sentence). The second approach assumes that free focus does not make by itself a semantic contribution. All it does is provided an informational and/or syntactic structure that comes into play when a so-called focus sensitive operator occurs in the sentence (e.g., an adverb of quantification). I believe that the data of the present paper confirm that, at least for Hungarian, the first approach is correct. The existence of the group denoter constraint on DPs in Focus, which, as I argue below, follows straightforwardly from a particular formulation of exhaustivity, in turn confirms the reality of exhaustivity itself. This is not to deny, though, that free focus in Hungarian also introduces a partition that focus sensitive items can rely on.

2.2. Retrospective: Exhaustivity for GQs over individuals

Let us first briefly consider an early attempt to formulate exhaustivity ("and no one/nothing else"). Szabolcsi (1983) proposed an interpretation for free focus in Hungarian, and Groenendijk–Stokhof (1984) for short answers to interrogatives in English. Groenendijk–Stokhof (1984, 419) note that the two formulations are logically equivalent but theirs is more transparent, so I will stick with it.

This definition of EXH takes a generalized quantifier Σ and returns its minimal element(s):

(17) Minimal element-EXH:
$$\lambda\Sigma\lambda P[\Sigma(P) \quad \& \quad \neg\exists P'[\Sigma(P') \quad \& \quad P \neq P' \quad \& \quad \forall x[P'(x) \rightarrow P(x)]]]$$

Does this notion of exhaustivity explain why only (set) group denoters are in Focus? Indeed, it turns out that (17) works well only with certain GQs. But what are they? They are the ones whose minimal elements are "characteristic" of them, e.g., *every man*, *no man*, *John and Bill*, *exactly five men*, but it misinterprets all GQs whose minimal elements are "uncharacteristic." E.g., EXH($[\![$*more than ten men*$]\!]$) comes out the same as EXH($[\![$*exactly eleven men*$]\!]$), and EXH($[\![$*fewer than ten men*$]\!]$) as EXH-($[\![$*no man*$]\!]$). Groenendijk–Stokhof address this problem and, wishing to provide a generally applicable definition, propose to remedy it by interpreting GQs with intersective determiners as follows:

(18) $\{X : \{G\} \subseteq X$, where G is a group of girls having n members$\}$

This proposal only derives certain "specific" readings, however, which may be acceptable for short answers but is not for Focus, which can be referentially dependent on wider scope quantifiers. Note also that whether or not we assume (18), the cases to which Minimal element-EXH applies correctly do not coincide with the GQs denoted by Type A DPs, therefore Minimal element-

EXH cannot be the answer to the question we are interested in. Prior to going further, let me note in passing that an initially quite plausible way to remedy (17) as a generally applicable definition would be to follow "Zwarts' Advice": If you get into trouble with minimal elements, switch to witness sets. The alternative definition is as follows:[8]

(19) Witness-EXH: $\lambda\Sigma\lambda P[\Sigma(P)\ \ \&\ \ P \subseteq SL(\Sigma)]$, where $SL(\Sigma)$ is the smallest set the generalized quantifier Σ lives on:
$\cap\{X : \forall P[P \in \Sigma \ \leftrightarrow\ (P \cap X) \in \Sigma]\}$

Let us see why this definition works where it does. Except for some principal filters, $SL(\Sigma)$ is the DP's common noun set. So, Witness-EXH(⟦Det *girls*⟧) says that Det girls have property P and only girls have property P. If ⟦Det *girls*⟧ is decreasing or non-monotonic, Det already sets an upper limit for girls, so we are fine. If it is innocently increasing like *more than six girls*, then no limit needs to be set for girls, so we are fine again. If Σ is a principal filter like the *six girls*, then its smallest live-on set $SL(\Sigma)$ may be smaller than the common noun set: it is the set consisting of the six individuals pointed at. This is exactly what we need. There is only one problem case: that of "non-innocent" increasing GQs, e.g., ⟦*two girls*⟧. However we define it, the application of Witness-EXH to it will give the same result as its application to ⟦*at least two girls*⟧. Thus this promising exercise eventually fails.[9]

2.3. Free focus: Presupposition versus assertion

I wish to argue now that the basic notion of exhaustivity in Szabolcsi (1981a,b, 1983) was misguided. Once it is amended, the question why only set (group) denoters express exhaustivity in Focus position is easily answered, too.

The problem with the notion proposed in my earlier work, technical details aside, is that it makes exhaustivity an asserted, rather than presupposed, part of meaning. The motivation for this decision came from the following type of contrasts:

[8] In referring to the smallest live-on set I assume that the universe is not (crucially) infinite.

[9] Construing *two girls* as a dynamic quantifier should help; this idea needs further work, however.

(20) *Nem [$_F$ Marit] veri János, hanem ...*
(a) ... *[$_F$ Katit].*
(b) ... *[$_F$ a porszívóval] játszik.*
(c) ... *[$_F$ az ajtó] csapódott be.*

'János is not beating MARI, but ...
(a) ... KATI.
(b) ... is fooling with the VACUUM CLEANER.
(c) ... the DOOR banged'

Clearly, the (b) and (c) continuations do not require that János beat anyone. Since these examples exhibit the same surface syntax as the ones discussed earlier, viz., have just DPs in Focus, I found it desirable to subsume them under the same interpretation rule. Now I believe that the cost of uniformity, i.e. the distortion of the interpretation of free focus, is too high, and these data should be given an independent account. Zsámboki (1995) proposes that the upward percolation of the focus feature gives rise to VP-focus and S-focus in the problematic cases.

Following Chomsky (1976), Kenesei (1986) proposed that free focus in Hungarian expresses, not exhaustive listing but, rather, exclusion by identification. Under this view, (21a) still requires that Mari and only Mari sleep, but does not assert all of this. In Kenesei's (1986, 149f.) formalization:

(21) (a) *[$_F$ Mari] alszik.*
Mari sleeps
(b) ιx[sleep(x)] = m

On the intended interpretation (21b) presupposes that there is a unique individual who is asleep, and asserts that this individual is Mari. I believe that this is essentially correct. Nevertheless, the formalism is too restrictive, since it cannot handle plurals. Notice that simply assuming that x can range over i-sums is not enough. For instance, [$_F$ *Mari és Kati*] *alszik* cannot be formalized as $\iota x[sleep(x)] = [m \oplus k]$, since this would not allow the correct distributive inference that Mari is asleep. One obvious way to correct this is as follows:

(22) EX-BY-ID: $\lambda z \lambda P[z = \iota x[P(x) \;\&\; \forall y[P(y) \rightarrow y \subseteq x]]]$

Van Leusen–Kálmán (1993), who discuss the pertinent presuppositions in fine detail, make the same proposal within their own framework:

(23) (a) Exhaustivity condition: for some discourse referents X, C, s.t. $X \leq_{c/o} C$, $c/o \models \mathrm{Max}(R, [\![X]\!]_{c/o}, [\![C]\!]_{c/o})$, where $\mathrm{Max}(P,U,V) \Leftrightarrow P(U) \ \& \ \forall U' \subseteq V[P(U') \Rightarrow U' \subseteq U]$

(b) Instruction to update context: $c/1 \models \mathrm{Max}(R', [\![F]\!]_{c/1}, [\![C]\!]_{c/1})$

Note, by the way, that if there is someone who sleeps then, assuming cumulativity, there is always a unique i-sum of those who sleep, so we are really only committed to an existential presupposition. It is identification that does the real work here.

For instance, interpreting two men as in (24), we get (25):

(24) *two men* $\Rightarrow \lambda P \exists g[\mathrm{man}(g) \ \& \ |g|=2 \ \& \ P(g)]$ where P is a vbl over sets of groups

(25) $[_F$ *Két ember*$]$ *fut* $\Rightarrow$
$\exists g[\mathrm{man}(g) \ \& \ |g|=2 \ \& \ g = \iota x[\mathrm{run}(x) \ \& \ \forall y[\mathrm{run}(y) \rightarrow y \subseteq x]]]$

These formulations easily explain why only set (group) denoters are input to exclusion by identification: the definition only works for such DPs.

Of course, it would be possible to define group readings for DPs not in Type A, but then we would make incorrect predictions for either collective readings or anaphora or both. Or, we might leave the interpretation of those DPs intact but go out of our way to construct a group as part of the definition of exhaustivity (see Bonomi–Casalegno 1993 (15) for such an exercise); but why would we want to do that if there are no data to motivate it?

3. *Csak* 'only'

3.1. Exhaustivity and association with Focus

Csak and *csupán*, both meaning 'only', can modify Type A noun phrases (26) and a subset of Type C ones: the non-monotonic or downward monotonic ones (27)–(29). The modification of Type B is excluded (30).[10] As for interpretation,

[10] In addition to attaching to the modified phrase, *csak* may also cliticize to the finite verb and associate with the Focus or Verbal Modifier. To my best knowledge, this does not make a difference. On the other hand, notice that phrases modified by *még* ... *is* 'even' occur strictly in Type B, which is excluded from Focus. This is of some interest since in English, *even* is assumed to be a focusing operator. I suspect that the phrase it associates with is merely intonationally prominent but not focused.

my judgment is that the same restriction applies here as with free focus: only Type A DPs get interpreted exhaustively. (On some variation in judgment, see the end of this subsection.)

(26) *Csak hat fiú ment el.*
only six boy went away
'Only six boys left (no one else did)'

(27) *Csak hat és tíz közötti fiú ment el.*
only six and ten between-six boy went away
*'Only between six and ten boys left (no one else did)'
'The number of boys who left is only between six and ten'

(28) *Csak hatnál kevesebb fiú ment el.*
only six-than fewer boy went away
*'Only fewer than six boys left (no one else did)'
'The number of boys who left was only fewer than six'

(29) **Csak hatnál több fiú ment el.*
only six-than more boy went away

(30) **Csak minden fiú / sok fiú elment / ment el.*
only every boy many boy away-went went away

Similar to (27) is *Csak hat fiú ment el* 'The number of boys who left was only (exactly) six' when *fiú* 'boy' is destressed. See note 7.

Let me first consider the issue of exhaustivity (the monotonicity contrast will be taken up in **3.3**). Does the restriction to set (group) denoters follow from the semantics of *csak*, too?

Csak-sentences are true under the same circumstances as sentences with free focus (well, almost—I come back to some differences below), but they are false under quite different ones. That is, the division of labor between presupposition and assertion is different in the two cases. Consider:

(31) (a) *Nem Mari ment el, hanem Kati (*is).*
not Mari went away but Kati too
'It is not Mari who left, but Kati (*too)'

(b) *Nemcsak Mari ment el, hanem Kati *(is).*
not-only Mari went way but Kati too
'Not only Mari left, but Kati *(too)'

In (31a), *Mari ment el* with free focus presupposes that someone left and asserts that it is Mari. In (31b), *Csak Mari ment el* presupposes (among other things) that Mari left and asserts that no one else did.

The correct interpretation of *csak* is in fact rather close to what I proposed (incorrectly) for free focus in earlier work, cf. **2.2.** This already indicates that the meaning of *csak* does not require the modified DP to be a set (group) denoter. Indeed, Bonomi–Casalegno (1993) propose an event-semantic explication of *only* that applies successfully to all DP denotations (once some of them are fixed a little) as well as to other focused categories (numeral, verb, argument pair, etc.):[11]

(32) If A,B are of type <event,t>, then only(<A,B>) = $\lambda e[B(e) \ \& \ \forall f[A^*(f) \rightarrow \exists g[B(g) \ \& \ f \subseteq E\ g]]]$, where A^* is the exist. closure of A, and A is obtained from B (roughly) by replacing the focus category of B by the appropriate skeleton (variable).

This says: The focusless version of the sentence, B (say, *Mary left*) is true, and for every event f of which A* (here: *Someone left*) is true, there is an event g, such that B is true of g and f is part of g. That is, every event of someone leaving is part of an event of Mary leaving. This could be improved by making B a presupposition; a useful and easy move.

What explains the set (group) denoter constraint, then? The fact that *csak*, just like English *only*, associates with Focus may be a good candidate. That is, the restrictor of the universal quantifier over events is defined by replacing the Focus phrase with a variable. Recall now that only Type A DPs occur in Focus *qua* DPs; Type C DPs can at best be pied piped to Focus, e.g., when e.g. their numeral is contrastive. It follows that the restrictor of the universal is only definable using a DP-skeleton when DP happens to be a set (group) denoter. In other words, this is more of a syntactic than a semantic effect.

I wish to note that some speakers of Hungarian also allow for an exhaustive interpretation of Type C DPs with csak. These speakers seem to allow

[11] I thank Sara Inclán, Manuel Español-Echevarria, and Matt Pearson for the observations.

defining the restrictor using a DP-skeleton even if DP got into Focus due to pied piping, not on its own right.

Crosslinguistically, there is an even more interesting contrast that calls for further research. Some speakers of English and Spanish judge that DP-modifier *only* and *solo* do not easily co-occur with Type C DPs, while VP-modifier only and solo do, and allow for an exhaustive interpretation:[12]

(33) (a) ??*I have only fewer than ten chairs.*
(b) *I only have fewer than ten chairs.*
'... and nothing else'

(34) (a) ??*Hay solo menos de diez sillas en la habitación.*
have only fewer than ten chairs in the room
(b) *Solo hay menos de diez sillas en la habitación.*
only have fewer than ten chairs in the room
'There are fewer than ten chairs in the room and nothing else'

3.2. Conservativity

Let me briefly comment on an important aspect of Bonomi–Casalegno's (1994) schema for *only*. *Only* is the traditional example of a non-conservative operator, in view of the following paraphrase:

(35) Only John and Bill are Frenchmen $\neq$ Only John and Bill are Frenchmen who are either John or Bill

The precise semantic contribution of the word *only* remains unanalyzed here. What happens now if it is made explicit? Bonomi–Casalegno effectively translate *only* as a universal whose restrictor is the focus frame, and on this construal *only* is fully conservative. Compare the following simplified versions:

(36) Only John and Bill are Frenchmen
'Every Frenchman is either John or Bill' = 'Every Frenchman is a Frenchman who is either John or Bill'

This suggests that conservativity may be way more pervasive than is generally assumed; it may be a general condition on the syntax/semantics inter-

[12] I thank Sara Inclán, Manuel Español-Echevarria, and Matt Pearson for the observations.

face. Thus, if surface syntax does not bear it out, it serves as a good reason for assuming a more abstract syntactic structure that does. In our case the resulting "tripartite structure" ONLY(FOCUS FRAME)(FOCUS) coincides with standard logical forms.

See Herburger (1993) for related insights based on noun phrase internal focusing.

3.3. Two further aspects of *csak*

To conclude the discussion of *csak*, I briefly point out two further aspects of its meaning.

Bonomi–Casalegno's (32) does not yet explain the infelicity of csak in the following text:

(37) *Senki sem szokott meglátogatni vacsoraidőben. De tegnap meghívtam Jánost és Marit, úgyhogy 8-kor (*csak) ők csöngettek az ajtómon.*
'Nobody tends to visit me at dinner time. But yesterday I invited John and Mary, so at 8 p.m. (*only) they rang my bell'

The problem is that *only* requires for it to be possible and in some sense expected for more people to have rung my bell, but the context does not license this presupposition; rather, on the contrary.

Adding such an ingredient to the characterization of *csak* will, *mutatis mutandis*, also explain the fact that csak can only modify the numeral part of a non-monotonic or downward monotonic Type C DP, cf. (26)–(27) and (28) above. With upward monotonic DPs, there cannot be "more":[13]

(38) *Csak hat és tíz közötti fiú ment el.*
'The boys who left were between six and ten; they could have been more but were not'

(39) *Csak hatnál kevesebb fiú ment el.*
'The boys who left were fewer than six; they could have been more but were not'

[13] For some reason such a scalar use of *only* is marginal in English. This does not affect the argument concerning the contrast in Hungarian.

(40) **Csak hatnál több fiú ment el.*
'The boys who left were more than six; they could have been more but were not'

Another aspect of csak that Bonomi–Casalegno's (32) does not predict is borne out by (41):

(41)	*Csak*	*Mari*	*és*	*Kati*	*emelte*	*fel*	*az*	*asztalt*	*(Juli nem).*
	only	Mari	and	Kati	lifted	up	the	table-acc	Juli not

'Only Mari and Kati lifted the table (Juli did not)'

Even if lifting is understood collectively, (41) does not mean that Mari and Kati acted by themselves, rather than as part of a bigger group. It may only mean that there was not any other independent table-lifter (individual or collective). Similarly, the negation of (41) would not mean that Mari and Kati were part of a bigger group; it would mean that there was at least one more independent lifter.

The reformulation of (32) to take care of these observations remains a task for further research.

References

Beghelli, F. 1993. A minimalist approach to quantifier scope. In: Proceedings of North East Linguistic Society 23.

Beghelli, F.–Ben-Shalom, D.–Szabolcsi, A. 1993. When do subjects and objects exhibit a branching reading? In: Proceedings of West Coast Conference in Formal Linguistics XII.

Ben-Shalom, D. 1993. Object wide scope and semantic trees. Proceedings of Semantics and Linguistic Theory 3.

Bonomi, A.–Casalegno, P. 1994. ONLY: Association with focus in event semantics. Natural Language Semantics 2.

Chomsky, N. 1976. Conditions on rules of grammar. Linguistic Analysis 2.

Dobrovie-Sorin, C. 1993. What does QR raise? In: The Syntax of Romanian. Mouton–de Gruyter.

É. Kiss, K. 1991. Logical structure in linguistic structure: the case of Hungarian. In: Huang, J.–May, R. (eds): Logical Structure and Linguistic Structure. MIT Press.

É. Kiss, K. 1994. Word order in simple sentences. In: Kiefer, F.–É. Kiss, K. (eds): The Syntactic Structure of Hungarian. Syntax and Semantics 27. Academic Press, New York.

Farkas, D. 1993. Modal anchoring and NP-scope. Ms. UC Santa Cruz.

Groenendijk, J.–Stokhof, M. 1984. The semantics of questions and the pragmatics of answers. Doctoral dissertation, University of Amsterdam.

Herburger, E. 1993. Focus on noun phrases. WCCFL 12.

Kamp, H.–Reyle, U. 1993. From Discourse to Logic. Kluwer, Dordrecht.

Kanazawa, M. 1994. Weak vs strong readings of donkey sentences and monotonicity inference in a dynamic setting. In: Linguistics and Philosophy 17: 109–59.

Kenesei, I. 1986. On the logic of Hungarian word order. In: Abraham, W.–de Meij, S. (eds): Topic, Focus, Configurationality. John Benjamins, Amsterdam.

van Leusen, N.–Kálmán, L. 1993. The semantics of free focus. ILLC Prepublications, University of Amsterdam.

Liu, F-h. 1990. Scope and dependency in English and Chinese. Ph.D. dissertation, UCLA.

Stowell, T.–Beghelli, F. 1994. The direction of quantifier movement. GLOW Colloquium, Vienna.

Szabolcsi, A. 1981a. The semantics of topic-focus articulation. In: Groenendijk, J. *et al.* (eds): Formal Methods in the Study of Language II. University of Amsterdam.

Szabolcsi, A. 1981b. Compositionality in focus. In: Folia Linguistica 15: 141–61.

Szabolcsi, A. 1983. Focussing properties, or the trap of first order. Theoretical Linguistics 10.

Szabolcsi, A. 1994a. Quantifiers in pair-list readings, and the non-uniformity of quantification. Proceedings of the Ninth Amsterdam Colloquium.

Szabolcsi, A. 1994b. Compute and look-up: two strategies of scope taking. Ms. UCLA and HAS.

Szabolcsi, A.–Zwarts, F. 1993. Weak islands and an algebraic semantics for scope taking. Natural Language Semantics 1.

Address of the author: Anna Szabolcsi
Department of Linguistics, UCLA
405 Hilgard Ave.
Los Angeles, CA 90024–1543
USA
and
Research Institute for Linguistics
Hungarian Academy of Sciences
Színház u. 5–7.
1250 Budapest, P.O. Box 19.
Hungary

Acta Linguistica Hungarica, Vol. 42 (3-4), pp. 189-207 (1994)

ANALOGY COMPREHENSION AS COMPLEXITY REDUCTION

BART VAN KLINK-LAMBÈR ROYAKKERS

1. Introduction

Metaphors are intriguing puzzles. How is it possible that language users understand an expression which joins two or more previously unconnected objects and which is, in most cases, literally false? The character of this process of interpretation is largely obscure. In an influential article Max Black (1976, 38 ff.) has suggested that in a metaphor like *A is B*, commonplaces that are generally associated with one object (e.g. wolf), are projected on another object (e.g. man). Appealing though this interaction view may seem, it does not tell exactly which associations are transferred, nor why. In the well-known wolf-case, it is apparent that only some of the associated commonplaces are relevant for the understanding of the metaphor (especially [agressive]) and others are not ([hairy], [with four legs], [with sharp teeth], etc.). A systematic explanation of this selection process is lacking.

In this paper we want to offer a logical reconstruction of analogy comprehension by means of a step-by-step model of complexity reduction. We focus on relatively new, 'living', metaphors in ordinary discourse, not on 'dead' metaphors, literal comparisons and metaphors in literary, scientific or other specific uses of language. We assume that primarily a communicative intention lies behind these metaphors—although they may appear improper or even shocking at first sight: with them a sender wants to convey to his or her audience a message that is meaningful, understandable and relevant in the given context.

The model presented here proceeds from a theory of similarity that was introduced in cognitive psychology by Amos Tversky (1977) and has been further developed within the framework of a theory of metaphor by Andrew Ortony (1984) and Eva Feder Kittay (1982; 1987). First, we will give a short overview of Tversky's theory of similarity and its application on metaphor theory by Ortony and Kittay. After that, our model is presented together with a logical representation of it.

1216-8076/94/$ 5.00

2. The creation of similarity

2.1. Tversky's Contrast Model

Tversky (1977) has developed a powerful theory of similarity for which he coined the name 'Contrast Model'. By means of this model Tversky tries to demonstrate that in a similarity judgment *Object A is like object B* the relative position and the context of the compared objects are important. In his view similarity has to be constructed as a 'set-theoretic process of feature matching'. Objects are not points in a coordinate space—as is often assumed—but sets of features. In a similarity judgment these sets of features are compared. Similarity is the outcome of the shared and distinctive features of the compared objects. According to the Contrast Model, similarity does not need to be symmetrical: in many cases the order in which the objects are placed, is not reversible, that is, not without changing the meaning of the judgment. Of course we can change the sequence in *Juliet is (like) the sun* to *The sun is (like) Juliet*, but this reversal does not express Romeo's intention very well. We also prefer to say that Madonna is (or pretends to be) similar to Marilyn Monroe than vice versa.

With respect to asymmetrical judgments of similarity, Tverksy makes a distinction between 'subject' and 'referent' which corresponds with, respectively, object A and object B in the comparison. Confronted with such a judgment, the receiver focuses on features of referent B and compares them subsequently with features of subject A. According to Tversky, we assume that the referent has the most salient features, and this is what a comparison justifies. In other words, we consider the referent to be a 'prototype' and the subject a 'variant' of the newly formed category to which A and B belong. Members of a category share prominent features of the prototypes and are connected with each other in the way that Ludwig Wittgenstein (1963, 32–3, Nr. 66–7) called 'family resemblance'. Because of the (relative) saliency of the prototype and the variant, we are able to make a comparative judgment of similarity in which the variant (or subject) is considered to be more similar to the prototype (or referent) than the other way round.

The crucial point of Tversky's model is that in an asymmetrical judgment of similarity a classifying statement is implicitly made: an object A is classed under a category C, of which object B is a representative or prototypical member (insofar as A shares features with B). When someone says: *In her latest video clip Madonna resembles Marilyn Monroe*, he or she possibly wants to classify the popular singer/actress in the category of attractive female stars that has Marilyn Monroe as a prototypical member. Within our existing conceptual scheme it is impossible to reverse the positions of both objects in the judgment (*Marilyn Monroe resembles Madonna*), because Madonna is not or not yet representative for this category. Because a comparison joins the subject and the referent, certain features of the referent are emphasized that need not have had that saliency before. Only in such a way the referent can serve as a prototype for the newly formed category. This category does not have to be unprecedented, and it is also possible that it combines objects which normally belong to two different categories.

In order to explain how the juxtaposition of two (or more) objects can stress some features and suppress others, we must give an account of the prominence of context in similarity judgments. When someone faces a set of objects, he clusters them—according to Tversky—in groupings such that (i) within each group the similarity among the objects is maximized and (ii) the group is maximally distinct from every other group; this is the so-called 'diagnosticity principle'. The given set of objects will emphasize and make salient certain features of these objects. In a set consisting of cherries and bananas on the one hand and lemons on the other, [sweet] is such a salient feature. This feature has diagnostic value: it indicates which objects belong to a certain group and which do not. Thus, the simple set of objects in our example can be divided in one group with sweet fruits (cherries and bananas) and another with non-sweet or sour fruits (lemons). On the contrary, [edible] is no distinctive or diagnostic feature, because all three objects can be eaten. Features can thus be characterized as diagnostic, when they determine objects as belonging to a certain classification.

When some objects are added to or removed from an existing set, the diagnostic value of the features which determine the clustering can change. A change in the diagnostic value of the features will also affect the perceived similarity between the objects. We might, for example, remove the lemons from our original set of objects and replace them by apples. A possible effect is that we decide to regroup the set in one group with cherries and apples and another with bananas. As a result of this regrouping the feature of taste [sweet] looses its diagnostic value, but the feature of form [round] is emphasized: by means of this feature we are able to differentiate between the two groups. Notice that the original classification in which cherries and bananas are similar is abandoned in the light of the new set of objects in favour of the new similarity judgment *Cherries are like apples*.

Thus, similarity is context-dependent: when changes in the grouping of objects occur, the diagnostic value of their features alters, leading to the perception of different similarities between these objects. Accordingly, we can infer that our perception of similarity between objects partly depends on the manner in which these objects are classified. Similarities do not only determine classifications, but are also determined by these classifications. "Thus", Tversky (1977, 344) concludes, "similarity has two faces: causal and derivative. It serves as the basis for the classification of objects, but is also influenced by the adopted classification."

2.2. Metaphor, simile and similarity

Following Tversky, Kittay (1982) considers metaphors and similes as classifying statements in which similarities between objects are revealed in a figurative manner. A metaphorical expression, which states the (always partial) identification of an object A with an object B, can thus be characterized as a special case of the similarity judgments that Tversky has described. Accordingly, the notions 'subject' and 'referent' in judgments of similarity coincide with the notions 'topic' and 'vehicle' that are used in metaphor theory. In a metaphorical identification the topic is filed under the (existing or novel) category of which the topic is the prototype. Because of this identification the common and jointly distinctive features of vehicle and topic acquire saliency.

It is noteworthy that two cognitive processes are carried out at once: firstly, features which the vehicle shares with the topic are stressed. As a result, the vehicle can serve as a prototype for the newly formed category in which the topic is the variant. Secondly, less prominent, 'forgotten' or even novel features of the topic come to the fore, and other features that are normally ascribed to it are pushed into the background. Here similarity shows its derivative (as opposed to its causal) face: a perception of similarity between objects is determined by the configuration in which these objects are contained, that is, by the given classification. Because metaphors and similes are able to emphasize novel or forgotten features of the topic, Black (1976, 37) is right to claim that they create similarity rather than record a previously existing similarity. This created similarity is, in turn, the basis for a novel, mostly transient, classification. The topic is put into a category to which it has normally—within our current conceptual order—no access.

However, Kittay points to an important difference between judgments which express a similarity in a figurative manner, and 'ordinary' judgments of similarity. In the case of ordinary similarity judgments, such as *A computer is an up-graded type-writer* or *Lime is like lemon*, we choose a specific referent in order to maximize the common features that distinguish the subject as well as the referent from other objects in the original set of objects. This referent is in general more prototypical than the subject, but they are already considered to be closely related objects. Comparisons to which ordinary statements of similarity give rise, take place in given, generally accepted categories. Ortony (1984, 191) calls this kind of comparison 'literal'. On the other hand, non-literal or figurative comparisons traverse existing classifications. The referent of a metaphorical statement (or vehicle) is chosen to maximize the differences of the subject (or topic) with other elements in the existing categories of which it is part. Actually, the more the vehicle differs from these elements, the better, because in this way the diagnostic value of the remaining features shared by topic and vehicle uniquely, increases substantially. Juliet has, like every other human being, very little in common with the sun, but by making the comparison the few similarities they share are emphasized, such as the central place both objects occupy in their respective universes. In the newly formed category the sun serves as a prototype and Juliet as a variant. As this example indicates, a figurative comparison is used when the usual classification(s) of the topic provides no, or no sufficiently apt, prototype for exactly those features which we want to attribute to the topic. According to Romeo, Juliet has no peer among us humans.

3. Analogy comprehension as a three-step process of complexity reduction

By means of Tversky's theory of similarity, together with additions from Kittay and Ortony, we are able to construct a general model for analogy comprehension. We use the notion 'analogy' in a neutral sense, for figurative as well as for literal judgments of similarity. As mentioned in the introduction, we are especially interested in uncommon metaphors in ordinary discourse. In this section the step-by-step process of complexity reduction will be discussed formally by means of set theory and logic. The approach to set theory and logic used here is a semantic one. For analytical reasons we distinguish three separate steps which overlap in reality (that is to say, psychologically). First of all, we discuss the corresponding semantics of features, objects, and categories.

3.1. The semantics of features, objects and categories

We define $\underline{F}$ as the set of atomic features. We use the letter $\underline{f}$ with possible marks for atomic features.

Definition 3.1. The language F_L of features is the smallest set given by the following rules:

1. F_L contains a given set $\underline{F}$ of atomic features.
   If $\alpha_1, \alpha_2 \in F_L$, then
2. the *choice* $\alpha_1 \mid \alpha_2 \in F_L$, pronounced as "$\alpha_1$ or α_2".
3. The *simultaneous* feature $\alpha_1 \& \alpha_2 \in F_L$, pronounced as "$\alpha_1$ together with α_2".
4. The empty feature $\underline{\emptyset}$.
5. The universal feature $\underline{u} \in F_L$.
6. The negated feature $\overline{\alpha} \in F_L$, pronounced as "not α".

In the following we shall use $\overline{\underline{F}}$ for the collection of negated atomic features: $\overline{\underline{F}} = \{\overline{\underline{f}} | \underline{f} \in \underline{F}\}$.

An atomic feature $\underline{f}$ (e.g. hairy) involves a (semantical) elementary feature f (hairy) possibly together with other elementary features (e.g. alive). The meaning of $\underline{f}$ only specifies the performance of a corresponding semantical feature f, but we are free to get any other set of elementary features simultaneously with f. Formally we give the semantics of features by means of sets that we will call *feature sets*. These feature sets denote 'packages' of (semantical) elementary features. We denote the set of (semantical) elementary features by F.

Definition 3.2. A *feature* set (f-set) is a not-empty subset of F.

Notation. The feature set possessed by an object O will be denoted as $[O]$. We use $O, O_1, O_2, ..., O', ...$ for objects. The set of all objects will be denoted by $\mathcal{O}$ and the set of all f-sets by $[\mathcal{O}]$.

For instance, let $F = \{f_1, f_2, f_3\}$, then the f-set $[O]$ consisting of the atomic features f_1 and f_2 is written as

$$[O] = \{f_1, f_2\},$$

and

$$[\mathcal{O}] = \{\{f_1\}, \{f_2\}, \{f_3\}, \{f_1, f_2\}, \{f_1, f_3\}, \{f_2, f_3\}, \{f_1, f_2, f_3\}\}.$$

Since our language of features contains a non-deterministic (choice) operator $|$, we have to consider sets of f-sets. The sets of objects are called *categories*.

Notation. We use $C, C_1, C_2, \ldots$ to denote categories and the sets of f-sets belonging to a category C will be denoted as $[C]$.

Our feature model $\mathcal{F}$ is now given by:

Definition 3.3. $\mathcal{F}$ is the collection of sets $[C]$ consisting of f-sets. The empty set $\emptyset$ is included in $\mathcal{F}$.

We shall use the operations $\cup$, $\cap$ and $^{-}$ on the domain $\mathcal{F}$ as the semantical counterparts of the syntactical operators $|, \&$ and h respectively, in our language F of features.

Definition 3.4. The semantic function $\| \ \| \in F_L \to \mathcal{F}$ is given by:

1. $\|\underline{f}\| = \{[O] | f \in [O]\}$
2. $\|\alpha_1 \mid \alpha_2\| = \|\alpha_1\| \cup \|\alpha_2\|$
3. $\|\alpha_1 \& \alpha_2\| = \|\alpha_1\| \cap \|\alpha_2\|$
4. $\|\overline{\alpha}\| = \|\alpha\|^{-}$
5. $\|\underline{\emptyset}\| = \emptyset$
6. $\|\underline{u}\| = \mathcal{O}$.

Remark. $\|\underline{f}\|$ expresses the meaning of feature (expression) $\underline{f}$: we specify the elementary feature f (simultaneous with some package of features). Thus only the feature f is determined. $\underline{u}$ stands for an arbitrary f-set.

Suppose α is the feature $\underline{f_1} \mid \underline{f_2}$, then $\|\alpha\|$ is the set of all f-sets consisting of the elementary feature f_1 or f_2. And so $\|\underline{f_1} \& \underline{f_2}\|$ is the set of all f-sets consisting of the elementary features f_1 and f_2.

Example. Let $F = \{f_1, f_2, f_3\}$. Then

- $\|\underline{f_1}\| = \{\{f_1\}, \{f_1, f_2\}, \{f_1, f_3\}, \{f_1, f_2, f_3\}\}$
- $\|\overline{\underline{f_1}}\| = \{\{f_2\}, \{f_3\}, \{f_2, f_3\}\}$
- $\|\underline{f_1} \& \underline{f_2}\| = \{\{f_1, f_2\}, \{f_1, f_2, f_3\}\}$.

In this section we also introduce an auxiliary notion that we shall use in the following:

Definition 3.5. We put $\alpha_1 =_{\mathcal{F}} \alpha_2$ iff $\|\alpha_1\| = \|\alpha_2\|$.

Proposition 3.6. $(\mathcal{F}, |, \&, h\ , \underline{\emptyset})$ is a Boolean algebra.

3.2. Step 1: The role of the vehicle

Our model is based on the following distinctions: manifest vs. latent features, general vs. specific features and ordinary vs. technical and unknown features. To begin with the first grouping: a feature of an object O is called 'manifest', if it has already been a feature with diagnostic value in at least one of the current categories associated with O (the object in question) within a linguistic system. That means that competent language users know that the feature in question has figured in prior classifications of this object O. It is important to stress that the category has to be current (either for scientists and experts or for ordinary people) instead of unusual or even obscure, because otherwise the feature does not have the cognitive prominence needed for a (relatively) uncomplicated understanding of analogy. Because of this prominence, a metaphor in ordinary discourse will probably refer to manifest features of the vehicle rather than to latent or unknown features. On the other hand, a latent feature has never been diagnostic in any current category to which the object belongs (but it may be so in the future) and is therefore less known.

The distinction between manifest and latent can be illustrated by the following example taken from Glucksberg–Keysar (1990, 6): *Sermons are like sleeping pills*. In this metaphor the vehicle 'sleeping pills', which belongs to the current categories of narcotics and medicines, according to common knowledge has certain features, such as [curative power], [calming effect], [sleep-inducing], [solid] and so on. The first two features of the vehicle are manifest, because the feature [curative power] is diagnostic in the category 'medicines' and [calming effect] in the category 'narcotics'. The other two features are latent, lacking diagnostic value: not all medicines or narcotics are sleep-inducing or solid.

Secondly, an object has a specific feature if at least one of the following two conditions is satisfied:

1. Only a few other objects in all current categories to which the object belongs also have this feature (i.e. an *intra*-categorical specific feature, for short intra-specific feature). With respect to medicines, for example, the feature [sleep-inducing] is more specific than the feature [curative power]. The last feature can be called general, at least with respect to the category of medicines (but not with respect to narcotics), because it is shared (or supposed to be shared) by all medicines.
2. Only a few objects in (at least one of) the related and current contrast sets share this feature (i.e. an *inter*-categorical specific feature, for short inter-specific feature). A contrast set is a category of objects that maintains a relation of opposition with (one of) the set(s) of objects (i) to which the original object (here, the vehicle) belongs; (ii) which objects share at least one manifest feature with the objects of the contrast set; and (iii) which objects are not members of the contrast set. In our example, a possibly current contrast set of narcotics is the category of pep pills, which both

have strong effects on our consciousness, but in a fundamentally different way (relaxing vs. activating). The category 'stones', however, is excluded as a contrast set by the requirement of relatedness: although they both share the manifest feature [material], stones are not commonly contrasted with narcotics. According to the selected contrast set, [calming effect] is more specific than [material].

As these examples indicate, the particularity of a feature is often a matter of degree. A feature of an object becomes less specific (and so more general) if more objects in all current categories of which O forms a part have this feature or in related and current contrast sets. It is very unlikely that an object will be used as vehicle for a metaphor on the ground of general features, because they are so common. More specific features, which are particular for a few objects, are better suited for metaphor production in ordinary discourse.

Finally, a feature is considered to be 'ordinary', if it is part of common sense knowledge that the object under consideration has this feature. If the feature is only known to (a group of) scientists or experts, then the feature is 'technical'. For instance, the feature [sleep-inducing] of the sleeping pills is ordinary, while their chemical structure is part of technical knowledge. An unknown feature does not form part of our cultural (technical nor ordinary) knowledge but may become a manifest or latent feature. A feature of which we are sure (perhaps falsely) that it cannot be attributed to an object, is not an unknown feature, but simply no feature of that object. As may be expected, ordinary features will be more important than technical and unknown features for our case.

Our central hypothesis is that generally the manifest, (intra- and inter-)-specific, and ordinary features of the vehicle *or* the intra-specific and ordinary features will be pertinent to the understanding of a metaphor in ordinary discourse. All latent, general, and technical features are thus ignored. This is what we call the first step in the process of complexity reduction. Only a part of the remaining set of manifest, specific, and ordinary features (in the first case) or intra-specific and ordinary features (in the second case) will figure in the new or already existing category of which the vehicle is the prototype and the topic is the variant, however. Which part is exactly selected will have to be decided in the next two steps.

After this first step the features of the vehicle 'sleeping-pills' in our example are reduced to [curative power], [calming effect] and [sleep-inducing] which are possibly relevant for the understanding of the metaphor. Other features, such as [solid], [material] and all chemical aspects, can therefore be ignored (see Fig. 1).

sleeping pills	manifest	latent	intra-specific	inter-specific	ordinary	technical
curative power	*			*	*	
sleep-inducing		*	*	*	*	
calming effect	*		*	*	*	
solid		*			*	
chemical structure		*	*	*		*
material		*			*	
etc.						

Fig. 1

The previous basic assumptions can be translated into mathematical formulas in the following way. Let $\mathcal{C}$ be the set of all current categories, $\mathcal{O}$ the set of all objects and F the set of all elementary features. We use $C, C_1, C_2, \ldots$ to denote categories, $O, O_1, O_2, \ldots$ to denote objects and $\alpha, \alpha_1, \alpha_2, \ldots$ to denote features. An object O possesses a collection of features, say $\alpha_1, \alpha_2, \ldots, \alpha_n$. We write $[O] = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, so $[O]$ is a set of features. A category C is a set of objects, say $O_1, O_2, \ldots, O_k$: $C = \{O_1, O_2, \ldots, O_k\}$. $O \in C$ if $[O]$ contains certain features essential for C, say $\alpha_1, \alpha_2, \ldots, \alpha_n$. This means $\{\alpha_1, \alpha_2, \ldots, \alpha_n\} = \cap_{O \in C}[O]$. It follows that

$$\text{if } C_i \subset C_k \text{ then } \cap_{O_j \in C_i} [O_j] \supseteq \cap_{O_j \in C_k} [O_i]. \tag{1}$$

For the collection of manifest features M_{C_i} in C_i this means that

$$M_{C_i} \subset \cap_{O_j \in C_i} [O_j]. \tag{2}$$

The manifest features of an object O_j (M_{O_j}) are the union of manifest features of the categories containing O_j:

$$M_{O_j} = \cup_{C_i} M_{C_i} \text{ with } O_j \in C_i. \tag{3}$$

From the previous formulas it follows that

$$\text{if } C_i \subset C_k \text{ then } M_{C_i} \supseteq M_{C_k}. \tag{4}$$

The other features of object O_j are the latent features (L_{O_j}):

$$L_{O_j} = [O_j] \setminus M_{O_j}. \tag{5}$$

For the collection of intra-specific features we need the following definitions:

- Let $\delta_{O_j}(\alpha)$ be the s-value of the feature $\alpha \in [O_j]$. The s-value is the degree in which the feature α for O_j is intra-specific.
- Let $\mathcal{C}_{O_j}$ be the set of all current categories to which O_j belongs ($\mathcal{C}_{O_j} \subseteq \mathcal{C}$).
- Let $\mathcal{O}_{O_j}$ be the set of all objects in $\mathcal{C}_{O_j}$ except O_j ($\mathcal{O}_{O_j} \subset \mathcal{O}$).
- Let $\mathcal{O}_{O_j}(\alpha)$ be the set of all objects in $\mathcal{O}_{O_j}$, containing the feature α and $\alpha \in [O_j]$.

In our model it follows that

$$|\mathcal{O}_{O_j}(\alpha_1)| \geq |\mathcal{O}_{O_j}(\alpha_2)| \text{ iff } \delta_{O_j}(\alpha_1) \leq \delta_{O_j}(\alpha_2), \tag{6}$$

with $|X|$ the number of elements of the set X.

The s-value $\delta_{O_j}(\alpha)$ is higher if less objects in $\mathcal{C}_{O_j}$ share the feature α. The s-value $\delta_{O_j}(\alpha)$ is lower if many objects in $\mathcal{C}_{O_j}$ share the feature α. In this way we determine the *relative* s-value: α_2 is more specific than α_1 for object O_j if $\delta_{O_j}(\alpha_2) > \delta_{O_j}(\alpha_1)$. Receivers have discretion when determining the s-value. For instance, we can define the s-value $\delta_{O_j}(\alpha)$ as

$$\delta_{O_j}(\alpha) = 1 - \frac{|\mathcal{O}_{O_j}(\alpha)|}{|\mathcal{O}_{O_j}|}. \tag{7}$$

The intra-specific features of an object O_j (S_{O_j}) are now the features of O_j with an s-value larger than a constant δ ($0 \leq \delta \leq 1$) and the other features of O_j are the general features (G_{O_j}). The value of δ is a matter of degree.

Example. Let $[O_1] = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7\}$ and $\mathcal{C}_{O_1} = \{C_1, C_2, C_3, C_4\}$, with $C_1 = \{O_1, O_2, O_3\}$, $C_2 = \{O_1, O_3, O_4, O_5\}$, $C_3 = \{O_1, O_5, O_6\}, C_4 = \{O_1, O_2, O_3, O_4, O_5, O_6\}$ and $[O_2] = \{\alpha_1, \alpha_5, \alpha_6, \alpha_7\}, [O_3] = \{\alpha_1, \alpha_2, \alpha_4, \alpha_5, \alpha_9\}$, $[O_4] = \{\alpha_1, \alpha_4, \alpha_6\}$, $[O_5] = \{\alpha_1, \alpha_2, \alpha_4, \alpha_8, \alpha_9\}$, $[O_6] = \{\alpha_1, \alpha_2, \alpha_9, \alpha_{10}\}$.

Suppose further that $M_{C_i} = \bigcap_{O_j \in C_i} O_j$ for $i = 1,2,3,4$. We can derive the following formulas:

1. $\mathcal{O}_{O_1} = \{O_2, O_3, O_4, O_5, O_6\}$, so $\mid \mathcal{O}_{O_1} \mid = 5$;
2. $\mathcal{O}_{O_1}(\alpha_1) = \{O_2, O_3, O_4, O_5, O_6\}$, $\mathcal{O}_{O_1}(\alpha_2) = \{O_3, O_5, O_6\}$, $\mathcal{O}_{O_1}(\alpha_3) = \emptyset$, etc.;
3. $\delta_{O_1}(\alpha_1) < \delta_{O_1}(\alpha_2) = \delta_{O_1}(\alpha_4) < \delta_{O_1}(\alpha_5) = \delta_{O_1}(\alpha_6) < \delta_{O_1}(\alpha_7) < \delta_{O_1}(\alpha_3)$;
4. $M_{C_1} = \{\alpha_1, \alpha_5\}$, $M_{C_2} = \{\alpha_1, \alpha_4\}$, $M_{C_3} = \{\alpha_1, \alpha_2\}$, $M_{C_4} = \{\alpha_1\}$;
5. $M_{O_1} = M_{C_1} \cup M_{C_2} \cup M_{C_3} \cup M_{C_4} = \{\alpha_1, \alpha_2, \alpha_4, \alpha_5\}$.

If we define the absolute s-value $\delta_{O_j}(\alpha)$ as mentioned above with $\delta = \frac{1}{2}$, then it will follow that $\delta_{O_1}(\alpha_1) = 0$, $\delta_{O_1}(\alpha_2) = \delta_{O_1}(\alpha_4) = \frac{2}{5}$, $\delta_{O_1}(\alpha_3) = 1$, $\delta_{O_1}(\alpha_5) = \delta_{O_1}(\alpha_6) = \frac{3}{5}$ and $\delta_{O_1}(\alpha_7) = \frac{4}{5}$, with the specific features $\alpha_3, \alpha_5, \alpha_6$ and α_7 of O_1 (S_{O_1}) and the general features α_1, α_2 and α_4 of O_1 (G_{O_1}). Now we can divide the features (f-set) of O_1 in 4 sets:

1. The set of manifest and intra-specific features: $\{\alpha_5\}$;
2. The set of latent and intra-specific features: $\{\alpha_3, \alpha_6, \alpha_7\}$;
3. The set of manifest and general features: $\{\alpha_1, \alpha_2, \alpha_4\}$;
4. The set of latent and general features: $\emptyset$.

For the collection of inter-specific features we need the following definitions:

- Let $\delta^*_{O_j}(\alpha)$ be the s-value of a feature $\alpha \in [O_j]$. The s-value is the degree in which the feature α for O_j is inter-specific.
- Let $\mathcal{C}^*_{O_j}$ be the set of all contrast sets of the categories C to which O_j belongs, with the following two conditions:
  1. $\forall_{C \in \mathcal{C}^*_{O_j}} O_j \notin C$;
  2. $\forall_{C \in \mathcal{C}^*_{O_j}} \forall_{O \in C} \exists_{\alpha \in M_{C_{O_j}}} \alpha \in [O]$.
- Let $\mathcal{O}^*_{O_j}$ be the set of all objects in $\mathcal{C}^*_{O_j}$. And $\mathcal{O}^*_{O_j}(\alpha)$ is the set of objects in $\mathcal{O}^*_{O_j}$, *not* containing the feature α.

The s-value $\delta^*_{O_j}(\alpha)$ will be defined as

$$\delta^*_{O_j}(\alpha) = 1 - \frac{|\mathcal{O}_{O_j}(\alpha)|}{|\mathcal{O}_{O_j}(\alpha)| + |\mathcal{O}^*_{O_j}(\alpha)|}. \tag{8}$$

The inter-specific features of an object O_j ($S^*_{O_j}$) are now the features of O_j with an s-value larger than a constant δ^* ($0 \leq \delta^* \leq 1$) and the other features of O_j are the general features ($G^*_{O_j}$).

Example. Let $[O_1] = \{\alpha_1, \alpha_2, \alpha_3\}$ and $C_{O_1} = \{C_1, C_2,\}$, with $C_1 = \{O_1, O_2, O_3\}$, $C_2 = \{O_1, O_3, O_4, O_5\}$ and $C^*_{O_1} = \{C_3, C_4, C_5\}$, $C_3 = \{O_6, O_7, O_8, O_9\}$, $C_4 = \{O_9, O_{10}, O_{11}\}$, $C_5 = \{O_{12}, O_{13}\}$; $\alpha_1 \in [O_i]$ for $i = 1, 2, \ldots, 13$; $\alpha_2 \in [O_i]$ for $i = 1, 2, 3$ and $\alpha_3 \in [O_i]$ for $i = 1, 3, 4, 5, 7, 8, 9, 13$. We can derive the following formulas:

1. $\mathcal{O}_{O_1} = \{O_2, O_3, O_4, O_5\}$, so $|\mathcal{O}_{O_1}| = 4$;
2. $\mathcal{O}_{O_1}(\alpha_1) = \{O_2, O_3, O_4, O_5\}$, so $|\mathcal{O}_{O_1}(\alpha_1)| = 4$;
3. $\mathcal{O}_{O_1}(\alpha_2) = \{O_2, O_3\}$ and $\mathcal{O}_{O_1}(\alpha_3) = \{O_3, O_4, O_5\}$;
4. $\mathcal{O}^*_{O_1} = \{O_6, O_7, \ldots, O_{13}\}$, so $|\mathcal{O}^*_{O_1}| = 8$;
5. $\mathcal{O}^*_{O_1}(\alpha_1) = \emptyset$, so $|\mathcal{O}^*_{O_1}(\alpha_1)| = 0$;
6. $\mathcal{O}^*_{O_1}(\alpha_2) = \{O_6, O_7, \ldots, O_{13}$, so $|\mathcal{O}^*_{O_1}(\alpha_2)| = 8$;
7. $\mathcal{O}^*_{O_1}(\alpha_3) = \{O_6, O_{10}, O_{11}, O_{12}\}$, so $|\mathcal{O}^*_{O_1}(\alpha_3)| = 4$;
8. $\delta^*_{O_1}(\alpha_1) = 1 - \frac{4}{4} = 0$, $\delta^*_{O_1}(\alpha_2) = 1 - \frac{3}{3+8} = \frac{5}{8}$ and $\delta^*_{O_1}(\alpha_3) = 1 - \frac{3}{3+4} = \frac{4}{7}$.

If the constant $\delta^* = \frac{1}{2}$, then α_2 and α_3 are inter-specific for the object O_1.

The set of ordinary features of an object O_j is called D_{O_j} and the set of technical features T_{O_j}. ($D_{O_j} \cap T_{O_j} = \emptyset$.) The unknown features are the other features of O_j: U_{O_j}. So $D_{O_j} \cup T_{O_j} \cup U_{O_j}$ are all the features of O_j.

For the understanding of a metaphor we are interested in $S_{O_j} \cap D_{O_j}$ or in $M_{O_j} \cap S^*_{O_j} \cap D_{O_j}$ from the object O_j (that is, the vehicle). This corresponds to the shaded parts in Figs. 2 and 3.

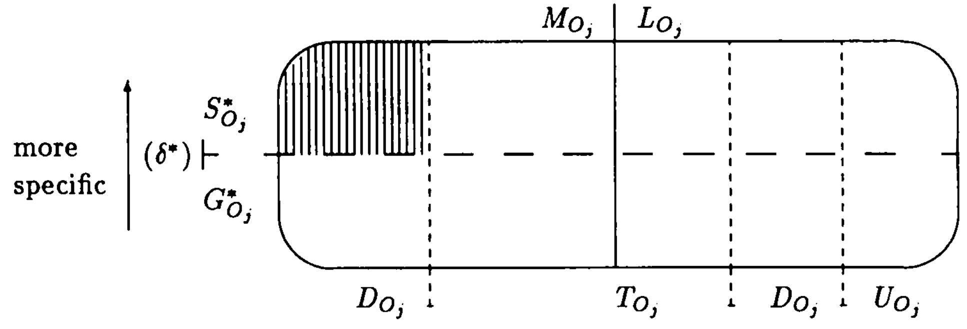


Fig. 2

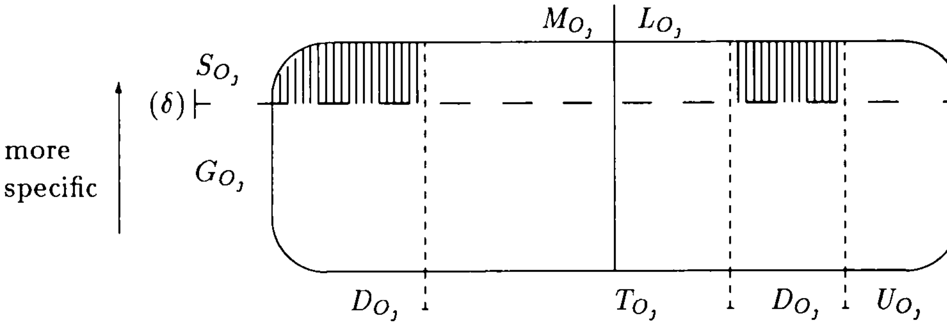


Fig. 3

Remark: Unknown features of a vehicle are latent by definition.

3.3. Step 2: The role of the topic

In step 2 the reduced set of (ordinary and intra-specific or ordinary, specific and manifest) features $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ of the vehicle O_j are interpreted in relation to the topic T. Only those features that can be applied meaningfully to the topic are selected. This is the reason behind the idea of a process of interaction between topic and vehicle. In this process, different interpretative strategies can be used: sometimes a feature can be applied almost literally to the topic (e.g. [calming effect] with respect to sermons), but sometimes the feature has to be interpreted figuratively to fit the topic (e.g. [sleep-inducing]). As a result, a collection of features for T is constructed:

$$\{I(\alpha_j) \mid 1 \le j \le n\}. \tag{9}$$

This is the collection of all possible interpretations of α_j ($I(\alpha_j)$) for all j belonging to topic T. The juxtaposition of topic and vehicle in the newly formed category has consequences for the following types of features of topic T:

1. Manifest features of the topic T: features that have already been a feature with diagnostic value in at least one of the current categories within a linguistic system (M_T).

2. Latent features: all known features of a topic T that are not manifest (L_T).

3. Unknown features: features that do not form or not yet form part of our cultural knowledge (U_T).
4. No features: all other features (N_T).

Figure 4 illustrates this for a topic T:

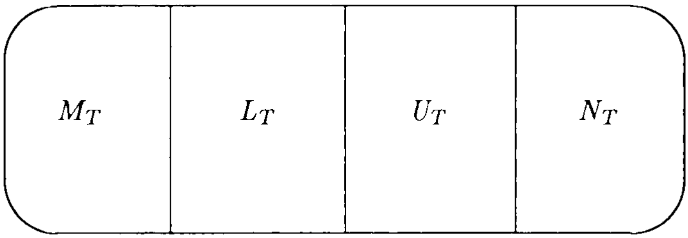


Fig. 4

The four possible consequences are:

1. A manifest feature α of T, *belonging to the reduced set* $\{I(\alpha_j) \mid 1 \leq j \leq n\}$ ($\alpha \in M_T \cap \{I(\alpha_j) \mid 1 \leq j \leq n\}$), becomes manifest again. In this case, the analogy is *literal*.
2. A latent feature α of T, belonging to the reduced set ($\alpha \in L_T \cap \{I(\alpha_j) \mid 1 \leq j \leq n\}$), becomes manifest. The analogy is *intelligible*, even without an explicit context.
3. An unknown feature α of T, belonging to the reduced set ($\alpha \in U_T \cap \{I(\alpha_j) \mid 1 \leq j \leq n\}$), becomes manifest too. The analogy is *difficult* to understand, or even *unintelligible*, without context information (cf. step 3).
4. A feature which was excluded from T, belonging to the reduced set ($\alpha \in N_T \cap \{I(\alpha_j) \mid 1 \leq j \leq n\}$), is now included. This means a fundamental reordering of our conceptual scheme(s).

This is illustrated in Fig. 5. The inner area is the reduced set $\{I(\alpha_j) \mid 1 \leq j \leq n\}$ and the numbers correspond to the numbers of the possible consequences.

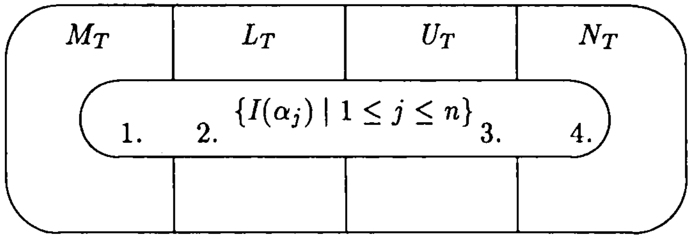


Fig. 5

After this second step the features of the vehicle in the example *Sermons are like sleeping-pills* are further reduced to two features: [calming effect] and [sleep-inducing]. The first feature can be applied almost literally to the topic: both sleeping-pills and sermons can be said to be calming, although they reach this effect in a different way. [Sleep-inducing], however, has probably to be understood in the sense of 'boring' when attributed to a sermon (although some people may actually fall asleep by listening to a priest). Both of these features are latent in the categories to which sermons belong (e.g. speeches, religious activities and so on), so the analogy can be considered figurative and also—even without an explicit context—intelligible. The feature [curative power] can be bracketed for the moment, because it is very unusual (though not unconceivable) to attribute healing effects—at least in the medical sense—to sermons.

sermons	manifest	latent	unknown	no feature
curative power			*	
sleep-inducing		*		
calming effect		*		

Fig. 6

3.4. Step 3: The role of the context

The last step has an important monitoring function: the interpreter checks whether the remaining set of features, applicable to topic T, makes sense within the (intra- and extra-linguistic) context in which the metaphor occurs (and which is responsible for the identification of an utterance as metaphorical). The interpretation of the metaphor must satisfy the requirement of coherence: it must be confirmed by the rest of the text (cf. Eco 1990, 64 ff.). Coherence is obtained, as Teun A. van Dijk and Walter Kintsch (1983, 190) show, by means of micro- and macro-strategies of interpretation.

Now there are three possibilities. In light of new data

1. the reduced set of features from step 2 can as a whole be applied to the topic;
2. the reduced set is even further reduced to fit the context;
3. or the reduced set is replaced by a new set of features, which will be subsequently applied to the topic.

Because of the last possibility, some features of the topic that were latent or unknown may become manifest in the newly formed category. This means that the collection of manifest features increases at the cost of the latent or unknown features. Further, new context information can make a metaphor meaningful that was unintelligible in isolation. After the three-step reduction an interpretation of the metaphor is reached.

In our example, no explicit context is given, but we can easily imagine one: a woman sits in a church and listens to a priest. As the priest goes on talking she gets tired of the lengthy and tiresome biblical exegesis and whispers to her neighbour "Sermons are like sleeping pills". In this situation, the outcome of the first two steps of complexity reduction is confirmed. Thus, the reader is prompted to take the metaphor as an expression about the boring character of sermons in general. Other, less stereotypical, contexts are possible of course, and they will generate new meanings, which also will by definition be less stereotypical (and therefore also more difficult to understand).

4. Final remarks

In this paper we have offered a model for analogy comprehension, especially suited for living metaphors in ordinary discourse. Logic supplies excellent tools for the description of the process of complexity reduction, and it is worthy of a thorough examination and futher development. An important question is whether our model has any relevance for other types of analogy as well. Although some further elaborations are necessary, we do think so. Our model explains when and why metaphors or similes will be generally considered trivial or dead. Literal analogies—as opposed to figurative ones—highlight features of the topic that were already manifest, specific, and ordinary within our cognitive system. They add no new information to our common sense knowledge and do not create novel or unusual categories. At best they emphasize some features of the topic that already have been highlighted in another category to which the topic belongs. But in most cases a comparison takes place within a literal analogy between the prototype and the variant of an existing category. Take for example, *Computers are type-writers*, *An encyclopedia is like a dictionary*, *Grape-fruits are similar to oranges*, and so on.

Besides, it is also possible to account for the experience of vagueness, obscurity or even unintelligibility to which some, especially literary, metaphors or similes give rise. How do we as non-professional readers make sense of T.S. Eliot's famous lines "Let us go then, you and I/When the evening is spread out against the sky/Like a patient etherised upon a table" ('The Love Song of J. Alfred Prufrock', 1982, 13)? And what exactly did Theodore Roosevelt mean by saying that a good political speech is "a poster, not an etching"? Analogies like these seem to be unintelligible, at least at first sight, because they address latent and/or technical features of the vehicle. Sometimes we have to dig very deep within our knowledge of the world to discover what is meant. Unless the sender supplies us with some more context information, the communication may fail. Fortunately, Roosevelt did so: according to him, in political speeches broad, bold themes that stand out from a distance are more important than fine details. And also fortunately, Eliot did not.

As Mary Hesse (1963) and others have shown, analogies as models may be essential to science. By exploring technical, latent, and even unknown features of the vehicle (e.g. a computer), we are able to learn more about the topic (e.g. the human brain). Thus, what may be a source of confusion in one situation, may become an important instrument of progress in another.

References

Black, M. 1976. Models and Metaphors: Studies in Language and Philosophy. Cornell University Press, Ithaca.

Dijk, T.A. van-Kintsch, W. 1983. Strategies of Discourse Comprehension. Academic Press, New York.

Eco, U. 1990. The Limits of Interpretation. Indiana University Press, Bloomington.

Eliot, T.S. 1982. The Complete Poems and Plays. Faber and Faber, London.

Glucksberg, S.-Keysar, B. 1990. Understanding metaphorical comparisons: beyond similarity. In: Psychological Review 97: 3-18.

Hesse, M.B. 1963. Models and Analogies in Science. Sheed and Ward, London.

Kittay, E.F. 1982. The creation of similarity: A discussion of metaphor in light of Tversky's theory of similarity. In: Philosophy of Science Association. Vol. 1: 394-405.

Kittay, E.F. 1987. Metaphor: Its Cognitive Force and Linguistic Structure. Clarendon Press, Oxford.

Ortony, A. 1984. The role of similarity in similes and metaphors. In: Ortony, D. (ed.): Metaphor and Thought, 186–201. Cambridge University Press, Cambridge.

Tversky, A. 1977. Features of similarity. In: Psychological Review 84: 327–52.

Wittgenstein, L. 1963. Philosophical Investigations 1. Basil Blackwell, Oxford.

Address of the authors: Bart van Klink and Lambèr Royakkers
Schoordijk Institute
Tilburg University
P.O. Box 90153
5000 LE Tilburg
The Netherlands
e-mail: B.M.J.vKlink@kub.nl, royakker@kub.nl

Acta Linguistica Hungarica, Vol. 42 (3-4), pp. 209-226 (1994)

SPECIFICITY*

PÉTER ÁBEL-MÁRTA MALECZKI

1. Introduction

The aim of this paper is to give an account of the many-sided character of specificity by a partial and dynamic semantic model.

In section **2** we sketch some observations from different approaches, and collect the most relevant properties used as starting points in our discussion. We establish the range and place of specificity along with emphasizing its gradual nature.

In section **3** we outline the framework that seems the most appropriate to grasp all these different characteristics. This is data semantics presented in Landman (1986).

In section **4** two interrelated sides of the specificity hierarchy are worked out under the headings of referential and property perspectives. The referential approach gives three definitions establishing the two extremes and the intermediate level. The property perspective refines this hierarchy. Thus we arrive at four definitions covering the wide range of phenomena subsumed under specificity. All these definitions are linguistically motivated as the conclusive examples illustrate.

In section **5** we summarize our results viewing them from wider perspective.

2. The notion of specificity

2.1. Views on specificity

In the relevant literature opinions differ concerning the exact nature of specificity. One crucial point of controversy is whether specificity should be treated as a semantic or as a pragmatic notion.

* We thank László Kálmán and László Pólos for valuable discussion which helped us improve our work as compared to the oral version presented at the symposium.

1216-8076/94/$ 5.00

Pragmatic accounts claim that specificity is a phenomenon that belongs to the use of language, therefore cannot be grasped truth-conditionally. Ludlow–Neale (1991) examine referents of indefinite NPs in various contexts, and they separate the definite, specific, referential, and purely quantificational uses of them. When characterizing the specific use of indefinites they heavily rely on the idea of "having an individual in mind". The very same idea is the starting point of the definitions given for specificity in Groenendijk–Stokhof (1980). However, they have a more extended notion of specificity, claiming that the specific/non-specific contrast can be applied not only to indefinites but to definites, numerals, singulars and plurals as well. They give definitions for the specific uses of the different determiners within the framework of epistemic pragmatics developed in their paper. The key notion of this framework is correctness, which is admittedly in accordance with the Gricean principles of cooperativeness. In the next sections we will examine the assumptions lying behind these pragmatic approaches in greater detail.

Semantic accounts of specificity also result in different analyses of the phenomenon. Fodor–Sag (1982) treat specificity in terms of scope relations; they say that an NP is specific if it has wide scope over at least one operator. However, the basic aim of their paper is not to investigate what the notion specificity can cover. Their crucial claim is that there are referential indefinites as opposed to quantificational ones, which is, in their opinion, due to a lexical ambiguity. We will offer another explanation for their data, considering their referential NPs to be specific (along with Enç 1991). Enç (1991) wants to treat specificity as an independent (from scope relations as well) semantic notion. She defines specificity in terms of discourse linkedness: non-specific NPs introduce new discourse referents, while the referents of specific NPs are from a previously introduced set of individuals. Definite NPs will automatically be specific, since they require identification with some previously introduced individuals. We will argue that Enç's proposal (also in its fully spelled out form) only partially covers the phenomena that can be subsumed under the heading specificity.

Any account of specificity has to take stand in this pragmatic/semantic controversy at least in its choice of framework and data to be accounted for. Since specificity seems to be also connected with some syntactic phenomena in some languages (see for example Reuland 1988; Enç 1991) we agree with the following argument from Diesing (1992, 80): "Since specificity does in fact correlate with a number of syntactic phenomena such as word order and case marking in a number of languages, it is reasonable to take the semantic approach to specificity, since syntactic effects might be unexpected under an account in terms of pragmatics."

However, we will offer a model-theoretic treatment for specificity as opposed to Diesing's approach which concentrates on the syntax/semantics interface.

2.2. Relevant properties of specificity

2.2.1. Graduality

Reexamining the empirical data presented in the earlier works and considering some other instances of specificity it seems to be crucial that specificity cannot properly be grasped in terms of a dichotomy: it is in fact a gradual notion. This is not a revolutionary idea at all. Ludlow–Neale (1991) separate cases of strong and weak specificity. In case of strongly specific use, the speaker wants the hearer to identify the referent of the indefinite NP, while this intention is not present in the case of weakly specific use. Our concept of specificity requires a different kind of graduality, more along the lines of Groenendijk–Stokhof (1980). They claim that the specific use of NPs requires a well-defined (set of) referent(s), while their non-specific use can be more or less close to the specific one, according to the number of the entities the NP can potentially refer to. So the epistemic pragmatic model they propose introduces grades in non-specific uses of indefinites, but the concept of graduality is not built into their definitions given for the specific uses of different kinds of NPs. We will propose that the notion of specificity should be built fundamentally on graduality.

Here are some examples for the gradual nature of specificity. First consider the following sentence and its three readings from Fodor–Sag (1982):

(1) *Each teacher thinks that a student of mine was called before the dean.*

(a) (each teacher: x) [x thinks that [(a student of mine: y) [y was called before the dean]]]

(b) (each teacher: x) [(a student of mine: y) [x thinks that [y was called before the dean]]]

(c) (a student of mine: y) [(each teacher: x) [x thinks that [y was called before the dean]]]

Reading (1a) is the least specific of the three: the teachers do not know which student was called before the dean, if any. Reading (1b) attributes different specific beliefs to each teacher. In (1c) each teacher has one and the same student in mind, therefore this is the most specific reading of the three. In this example the same sentence has three readings with respect to the degrees of the embedded subject's specificity. Another conspicuous instance of the gradual nature of specificity is when sentences share only their predicates, and different subjects show different levels of specificity explicitly. Examples illustrating this can also be given from Fodor–Sag (1982):

(2) *A student that Betty used to know in Arkansas cheated on the exam.*

(3) *A friend of mine cheated on the exam.*

(4) *Someone cheated on the exam.*

The lengthy description of (2) provides a number of properties of the cheater making it quite specific. The subject of (3) is less specific than (2), still it is obviously more specific than the vague *someone* in (4).

2.2.2. The range of specificity

Another relevant point which is to be clarified when examining specificity is that whether this phenomenon is restricted only to indefinite NPs. In this respect we agree with Groenendijk–Stokhof (1980), who claim that specificity concerns not only indefinites but all the other NPs. We give here only two examples to the extended range of specificity. The examples anticipate our claim that the partly overlapping observations occurring in different approaches can all be traced back to the fact that the more specific an NP is the easier is to identify its referent(s).

(5) *Two boys are playing in the garden.*

(6) *Two ugly loud boys are playing in the garden.*

(7) *Every girl is silly.*

(8) *Every girl in our class is silly.*

(9) *Every red-haired girl in our class is silly.*

Examples (5–6) and (7–9) show that NPs containing numerals or universal quantifier can be more or less specific as well. Different levels of specificity arise in the same way as in examples (2)–(4) above.

2.2.3. The place of specificity

The moral of the story so far is that specificity is a slippery notion because of its many-sided character. In the previous sections we have demonstrated that specificity cannot be treated as a binary notion and cannot be described satisfactorily if we restrict our attention to one group of determiners (i.e. to the weak ones). If we try to collect the factors influencing (or reflecting) the specificity of an NP, we find that these are several and seemingly unrelated, including at least the following:

i) the determiner of the noun

ii) the verb whose argument is the NP at issue; specifically:
  a) some lexically given features of the verb (intensionality, perfectivity, etc.)
  b) some properties of the sentence mainly expressed by the verb in several languages (person, number, time and modality)

iii) word order

iv) any explicitly given property of the NP expressed by:
  a) linguistic means (adjectives, relative clauses, etc.)
  b) non-linguistic means (deixis, shared background knowledge, etc.)

In our argumentation we will use (iv) as a starting point. It is the information in (ivb) which might give rise to the conclusion that specificity is a pragmatic, not a semantic notion. However, the answer to the question, what belongs to semantics and what to pragmatics, also depends on the (semantic) model we choose. Pragmatic accounts usually argue that specificity depends on what the speaker has in mind when uttering an NP. We certainly admit that the mental state of the speaker cannot belong to semantics. But we claim that specificity cannot be grasped only by referring to the speaker's mental state. Taking into account the hearer's mental state cannot help us either. What we need to do is to abstract away from the speaker's and the hearer's mental states and find the shared assumptions present in both of them. We will assume that specificity belongs to a level supposedly shared by both communicating parties. This abstract level of shared assumptions is the so-called intersubjective level as described in Landman (1986). Sketching the main ideas of Landman's partial model takes us to the next section; for the time being we want only to emphasize that in our opinion specificity can be placed best on a level where some supposed agreement is present between the speaker and the hearer. This level is that of semantics in Landman (1986), but the truth-conditions are based on information exchange. As a matter of fact, this is exactly what we need for grasping specificity.

As a starting point, we assume that an NP is wholly specific iff it is evident to both the speaker and the hearer that there is a certain entity they speak of. This entity may be present only very partially in the given information state, but if the NP referring to it is specific, the communicating parties mutually suppose that it is well-definable at least in some information state.

3. The theoretical frame

The discussion so far points to the conclusion that specificity, being a gradual notion connected to information exchange, cannot be given by a traditional semantic model. As we consider specificity a dynamic feature based on the (supposedly) available properties of the entity we refer to, we have chosen the partial and dynamic model described in Landman (1986). This model seems especially suited to our purpose as it makes a clear-cut distinction between the objects of the world and the objects of the information exchange. The basics of "Pegs and Alecs" (see in Landman 1986) are the following.

Classical models and classical identity conditions cannot be truly partial, by their very nature. However, we need a really partial device to model some phenomena of natural languages, especially identity conditions for linguistic objects (see the Hesperus-Phosphorus problem). For this purpose it is not enough to use partial models of the world. Partiality must be present not (only) at the level of the objects existing in the world independently of our mental states. Operating with partiality on the subjective level does not offer a solution to linguistic problems either. The most appropriate module for grasping natural language semantics is a level called intersubjective, which is to be distinguished sharply both from the mental states of the communicating parties (the subjective level) and the worlds of real objects. This intersubjective level is where information exchange takes place; the notions of this level are accessible for communicating parties through the basic assumption that there is a world with objects existing in their own right. Not knowing the world exactly is not a reason for giving up talking about it; however partial our information about the objects of the supposed real world is, we exchange information about them. Thus the basis of our communication is that we agree tacitly on the existence of objects belonging to the real world in spite of our very defective knowledge of them.

On the intersubjective level the counterparts of real world objects are partial objects, so-called pegs. **Pegs** are approximations of real world objects; the name "peg" indicates that they are like real pegs designed for holding clothes; and since "clothes" of objects are their properties, our pegs hold properties. Instead of giving objects as basic entities, the model will contain a set of basic properties, and objects (pegs) will be defined only through the properties attributed to them. Thus a peg is discernible only if it has some properties distinguishing it from other pegs, just like real pegs might be characterized by the clothes hung on them. Our cloak room attendants will be information states, it is their duty to hang properties on pegs.

An information state is a set of simple propositions (called facts) about pegs. Pegs do not **have** properties; properties are only attributed to them by an information state, by virtue of the facts the information state contains. Consequently, if an information state has two simple propositions about a (supposed) real world object, then this information state gives us a peg with exactly those two properties.

Facts constitute an **information system** defined in Landman (1986) in the following way.

An information system S is a triple $<P, \leq, 0>$, where:

1. P is a nonempty set of propositions
2. $\leq$, the relation of information containment, is a partial order on P
3. $0 \in P$ and for all $p \in P$: $0 \leq p$ (0 is the impossible fact)
4. if p, q $\in$ P then $p\&q \in P$, where p&q is the largest proposition such that $p\&q \leq p$ and $p\&q \leq q$.

Propositions are incompatible if the combination of their information is impossible, i.e. if $p\&q=0$.

As it can be seen from this definition, an information system is a meet semilattice, whose elements are any kind of propositions (so it will contain also incompatible facts). An information state, however, cannot contain incompatible pieces of information. An information state is one of those parts of an information system that contain only compatible propositions, and do not contain the impossible fact. (Thus it is a proper filter on the meet semilattice.)

A **model** is a structure $M=<S, D, \emptyset, f>$, where S is an information system as defined above, f is an assignment function, $\emptyset$ is a set of properties assigned to the predicates of the language by f, D is the domain of objects assigned to individual constants by f. As we identified objects with their properties, the elements of D will be defined by the information states, that is, by proper filters of the information system S.

The notion of **indiscernibility** will be important for our definitions to be given for the different degrees of specificity. Landman defines this notion with the aid of discernability in the following way:

A peg d_1 is discernible from d_2 on the basis of an information state s iff there is a property P such that: either $P(d_1) \in s$ while $P(d_2)$ is incompatible with some fact $f \in s$

or $P(d_2) \in s$ while $P(d_1)$ is incompatible with some fact $f \in s$.

Pegs d_1 and d_2 are indiscernible on the basis of s iff they are not discernible on the basis of s.

A **variable** is interpreted as a special peg: for awhile it is an indiscernible approximation of some pegs on the basis of some information state(s). Thus it can play the role of different pegs, at least during some information states, just like an actor (e.g., Alec Guiness) can play the role of every victim in a film. That is why Landman calls a peg which serves as interpretation for a variable an **alec**.

If the information system is a partially ordered set, and an information state is a proper filter on it, information states also can be partially ordered. We call the information state s' the proper extension of s (and write s < s') iff s' contains all the propositions s contains and some more. By extending the information states we might reach in principle at an information state w of which no proper extension exists, to which no more compatible pieces of information can be added. This is the total information state, which is the best approximation of the real world; actually, it can be considered as *being* the real world. Thus the real world can be considered a regulative idea or standard at the intersubjective level (see a more detailed discussion in Landman 1986).

For our purposes the most significant extension of an information state s will be what we call the **minimal extension of** s. We define it in the following way:

The information state s' is a minimal extension of s iff s < s', and there is no s" such that s < s" < s'.

4. Two perspectives of specificity

Now we are in a position to begin to speak about specificity. Our starting point will be the following generalization based on the argumentation in section **2**:

(10) An NP is the more specific the more properties of the entity (entities) it denotes are agreed upon by the speaker and the hearer(s).

Expressing this idea in Landman's system, specificity will certainly be some property of pegs, that is, a property about properties (recall that pegs can only be identified by their properties). In terms of pegs and information states (10) can be given as (11):

(11) A peg is more specific in an information state s than in s' iff s attributes more properties to it than s'.

Let us take a closer look at what it means "to attribute" some property to a peg. The simplest case is when we have a proposition in the information state directly expressing that property. For instance, uttering sentence (12) creates an information state with two simple propositions (shown in (13)):

(12) *A girl came in.*

(13) d is a girl
d came in.

The interpretation of d is that it is a peg with just the properties the proposition it occurs in attribute to it. If we add to our information state more and more properties, our peg becomes more and more specific. This approach might be called the "property" perspective.

However, if we take a look at this specification process from another point of view, we might gain further important insights. The more properties a peg has, the narrower its referential possibilities are: this other perspective we call the "referential" perspective of specificity. A peg can be considered specific according to its referential properties predictable on the basis of a given information state s. What might or must happen to our peg in the extensions of an information state s can be definitive for establishing the degree of specificity. A peg, by its very nature, can approximate some (more or less fuzzy) set of individuals or also a well-identifiable individual in the proper extensions of s.

This can be expressed by the determiner of the noun, by adjectives, or other kind of information attributing some property to the peg playing the role of the denotation of the noun phrase (see also the list given in **2.2.3**). The obvious example is the definite article; from its meaning it follows that there is exactly one relevant individual with the given properties. Another well-known example is the Turkish accusative morpheme, which can indicate that the individual referred to is a member of a previously given set (see Enç 1991).

A more detailed discussion and examples follow in the next section.

4.1. The referential perspective of specificity

The highest grade on the referential scale introduced above is when exactly one individual is identified, either explicitly or by means of presuppositions. The latter happens when using NPs with definite article or proper names.

Consider the meaning of the definite article. What is certain about a (singular) definite NP is that it refers uniquely. This can be expressed in the theoretical frame we have chosen in the following way. Take for example the information state s which contains the fact corresponding to a sentence containing the definite NP *the camel*. In this information state we will attribute the properties of a camel to a peg, plus we have the presupposition that there are some more properties which distinguish this camel from all the other ones. This presupposition can be expressed in our framework by the condition that in the minimal extensions of the information state s there is no more than one peg having the properties of a camel. We formulate this condition by stating that all pegs having the properties of a camel will be indiscernible on the basis of the minimal extensions of s.

Thus the truth-conditions of a proposition containing a definite NP can be expressed in the following way:

(14) If α is a nominal, β is a verbal phrase, then s $\models$ 'the $\alpha\,\beta$' iff there is a peg d such that s $\models \alpha(d)\ \&\ \beta(d)$ and for every minimal s' extension of s it holds that for every d' such that s' $\models \alpha(d')\ \&\ \beta(d')$, d' is indiscernible from d.

We agree with Groenendijk and Stokhof (1980) in that the number of possible referents is definitive when examining specificity. In fact, we will build our definitions exactly on these properties in this section.

Our starting point is that one well-distinguishable extreme for specificity is when the number of the possible referents (on the basis of a given information state) is no more than one. This is the strongest kind of specificity, at least from our current point of view. Thus proper names and (singular) definites are always specific in our sense (and in whatever sense). There is nothing revolutionary in this claim (see Enç 1991). However, while in Enç (1991) the specificity of the definite NP follows from the definition given for specificity, we reason in the opposite way: grasping definiteness in the way given in (14) makes it possible to define the highest extreme of specificity along the same lines.

(15) A peg d is strongly specific on the basis of an information state s iff d is such that s $\models P_1(d) \& \ldots \& P_n(d)$, and for every minimal s' extension of s it holds that for every d' such that s'$\models P_1(d') \& \ldots \& P_n(d')$, d' is indiscernible from d.

Comparing the definitions given in (14) and (15), it is clear that definiteness will be a subcase of strong specificity. The definition in (15) leaves open the possibility to be strongly specific purely on the basis of explicitly mentioned properties.

Thus also an indefinite NP can be strongly specific if we list enough properties to identify one individual. For example, it seems to be a well-motivated suspicion that the indefinite object NP in (16) refers to no more than one individual:

(16) Yesterday at 8:31 pm I met a green-haired and red-eyed girl whistling a melody from The Cats and wearing a large black hat and a yellow bathing-suite with white stripes.

The other end of the scale is the non-specific extreme, when there is no agreement even on the **existence** of the object referred to. This may happen in case of NPs which are arguments of intensional verbs. It is well known that intensional verbal phrases containing indefinites have two readings: beside a possible specific reading they also have a non-specific one. Thus in intensional contexts the possibility might arise that there is no agreement between the speaker and the hearer even on the existence of the object. This is what we consider the least specific (that is, non-specific) interpretation of a noun-phrase.

Exemplifying this idea by means of pegs and information states, take the sentence *Hugh is seeking a camel.* In the information state s we arrive at uttering this sentence, we attribute the properties of a camel to an alec x. Moreover, this alec also has the property of being sought by Hugh. By virtue of the latter property (the meaning of the intensional verb) it might turn out that Hugh cannot find any camel. In that extension of the information state s there only will be pegs with incompatible properties given for the alec x in s. Therefore every peg will be discernible from x, that is, there will be no peg looking like a camel.

(17) A peg x is non-specific on the basis of an information state s iff it is an alec with respect to the properties $P_1 \ldots P_n$ in s, and there is a minimal s' extension of s such that there is no peg d such that d is indiscernible from x.

Strongly non-specific readings may also arise when the sentence is in the non-indicative. Take for instance the imperative. When you turn to me with the request *give me a pencil*, you suppose that I have one, but there is no (presupposed or whatever) **agreement** between us on the **existence** of a pencil on me. Of course you might have seen my pencils earlier, and in this case *a pencil* can be specific in the sense that it belongs to a previously given set, but non-linguistic context is always able to specify any linguistically non-specific object.[1]

Between these two extremes of specificity are all the other referential possibilities: there is agreement on the existence of some individual, but this agreement is vague enough to leave open the possibility to refer to different individuals. The possible referents may be more or less specified according to the extent the set they belong to is narrowed down. We call this intermediate level vaguely specific, since it covers several degrees of specificity. Distinctions between these degrees cannot be based upon the purely referential approach, though it is certain that the less the number of the possible referents the more specific the NP is. However, pegs cannot be counted, so we leave the discussion of these differences to the next section.

To define the notion of vaguely specific, we only have to keep in mind that those NPs refer vaguely specificly which are neither strongly specific nor non-specific. As opposed to non-specific reference, there will be agreement on the existence of the object referred to in every extension of the given information state s. On the other hand, contrary to strong specificity, there is no

[1] The context can fix what the preferred reading actually is in the case of intensional verbs as well. If I utter the sentence *I am seeking a pencil* while trying to find one on my desk or in some usual place where a pencil can occur (that is, in a neutral place), then the possibility to interpret the NP non-specificly is much more probable as compared to the situation in which the same utterance sounds from under the bed (in a very unusual, that is, highly specific context).

agreement on exactly which individual is referred to. This can be expressed in the framework we work with by saying that some minimal extensions of s might make distinctions between the pegs having the properties s attributes to them. Thus the pegs indiscernible on the basis of s might become discernible on the basis of s' by having incompatible properties.

(18) A peg d is vaguely specific on the basis of an information state s iff d is such that $s \models P_1(d) \& \ldots \& P_n(d)$, and there will be some minimal s' extension of s where for some d' such that $s' \models P_1(d') \& \ldots \& P_n(d')$, d' is discernible from d.

On the basis of this definition, all the following examples contain vaguely specific NPs.

(19) *There are camels.*

(20) *I have found a camel.*

(21) *I have found one of the camels.*

Each peg having the properties of a camel in s corresponding to (19)–(21) can be attributed different properties in the extensions of s. That is, a camel can be further specified as white or brown, having one, two or three humps, etc. Thus we can specify these pegs further in several different ways. This possibility for "branching" reference as opposed to the unique reference of strongly specific NPs is the only feature necessarily shared by vaguely specific NPs.

It is evident that although we collapsed the NPs in these examples under the heading vaguely specific, they differ significantly with respect to their specificity. This forces us to take into consideration another point of view, in accordance with our intuition that these NPs represent different degrees of specificity.

4.2. The property perspective of specificity

In the previous section we defined three distinct areas of specificity. Three is certainly more than two, thus we replaced the binary specific/non-specific distinction with a "trinary" notion. However, the really intriguing part of the scale is the intermediate level placed between the two well-defined extremes. We labelled it as vaguely specific, so even the name itself indicates that we are badly in need of some further specification. Focussing on graduality, vaguely specific pegs also have to be compared with respect to their specificity.

Let us return to the other perspective offered in the beginning of this section, and try to keep track of a peg through information exchange. If we extend an information state s by attributing some property to a peg, it will certainly be more specific than it was in s. This was already expressed in (11), repeated here as (22):

(22) A peg is more specific in an information state s than in s' iff s attributes more properties to it than s'.

Let us consider the following sentences:

(23) *Hugh has found a camel.*

(24) *Hugh has found a camel with three humps.*

(25) *Hugh has found a one-eyed camel with three humps.*

Assume we are in an information state s corresponding to (24), and compare it to s' corresponding (23). The peg which has the properties of being a camel and having three humps is certainly more specific (can refer to fewer real world entities) than the one only specified as a camel.[2] This difference in specificity can be grasped by the observation that s is a proper extension of s', and s attributes more properties to the peg indiscernible from our relevant peg in s'. The same reasoning applies to the relationship between (25) and (24). Consequently, the specificity of two indiscernible pegs can be compared according to the definition in (26), which is a more precise version of (22):

(26) On the basis of an information state s a peg d is more specific than a peg d' on the basis of an information state s' iff s is a minimal extension of s' and if it holds that $s' \models F_1(d') \& \ldots \& F_n(d')$, then there is a P property such that $s' \not\models P(d')$ and $s \models P(d)$.

As we said in section **2.2.2**, the range of specificity is not restricted to indefinite noun phrases. A universally quantified NP can also be more or less specific, according to the extent the context narrows down the class of the relevant

[2] Provided that we remain in the same world.

individuals. Let us refer back to examples (7)–(9), containing the universal quantifier. The specificity hierarchy present in them can be grasped in exactly the same way as in (23)–(25).

Of course it might happen that so many properties are given that the universal quantifier will denote exactly one individual, becoming strongly specific in this way. By adding more and more properties to the peg corresponding to a universally quantified NP, we can ensure that our peg becomes more and more specific through extending information states, up to referring uniquely. Notice that our specificity definition given in (15) does not refer to the number of real world entities. Consequently, a universally quantified NP does not become strongly specific only by virtue of considering its denotation in a world containing exactly one individual in the extension of the quantified noun.

As we have seen in this section, the degree of specificity basically depends on the number of properties upon which there is agreement on the intersubjective level. However, it is not necessary that all these properties are expressed explicitly. It is enough if they are present by presuppositions. Consider example (21) as opposed to (20): the complex determiner *one of the* presupposes that the camel we refer to belongs to a previously mentioned set. Therefore, it has some more properties (those characteristic of the set) not explicitly specified in the given information state. Although both object NPs are vaguely specific by their referential properties, on the basis of the definition given in (26) the NP in (21) is more specific than the one in (20). Notice that this is the way we account for the specificity distinction Enç's definition covers.

An interesting Hungarian counterpart of the above examples for the presupposed agreement on a previously given set are perfective verbal prefixes. Verbal prefixes are connected lexically to (some) verbs, and they can change the meaning of the verb in very different ways. Perfectivity is also expressed by some of these prefixes in Hungarian. Compare the following examples:

(27) *Írtam egy levelet.*
wrote-1sg a letter-acc
'I have written a letter.'

(28) *Megírtam egy levelet.*
perf-wrote-1sg a letter-acc
'I have written one of the letters.'

Perfective verbs, as opposed to their non-prefixed counterparts, always presuppose some, but not all distinguishing properties of the entity their indefinite argument refers to. Perfectivity presupposes just as many distinguishing properties of the relevant entity that it can be considered an element of a previously identified set. Thus specification through perfectivity also arises by attributing more properties to a peg than in a previous information state.

All the examples mentioned so far in this section are vaguely specific from the referential perspective. However, further specification through properties is possible within strongly specific and non-specific NPs as well. We can consider the definite NP in (30) more specific than in (29) on the basis of the same argumentation we have given so far:

(29) *The camel is having breakfast in the kitchen.*

(30) *The white camel that we borrowed from an Arab yesterday is having breakfast in the kitchen.*

In intensional contexts non-specific NPs can be more or less specified according to the number of their explicitly given properties as well:

(31) *I dreamed about a white camel.*

(32) *I dreamed about a white, benevolent camel that gave me a cup of ice-cream.*

5. Summary

As we mentioned in **2.2.3**, there are various linguistic and extralinguistic means to express different degrees of specificity.

Some of them are exclusively connected to one well-definable kind of specificity. The question arises what happens if these seemingly incompatible factors interact in determining the specificity of one and the same peg.

As we saw in **4.1**, the definite article is characteristic of strongly specific NPs, while non-specific reference is connected to intensionality. Let us take an example containing a definite NP in multiply intensional context:

(33) *Leonora believes that Max is searching for the fictitious unicorn.*

The strongly specific definite NP makes impossible the non-specific reading in the intensional context. Recall that intensional contexts are ambiguous in that they leave open the possibility for an NP to become more or less specific.

This enables the definite determiner to exert its strongly specific effect on the NP. As example (30) shows this is true even in multiply intensional contexts. The definite determiner carries the presupposition in this context that there is some agreement on (some properties of) a certain imagined unicorn between the communicating parties.

Another intriguing case is when the differences in specificity seemingly arise from different scopes. Consider our example (1) taken from Fodor–Sag (1982), repeated here as (34):

(34) *Each teacher thinks that a student of mine was called before the dean.*

(a) (each teacher: x) [x thinks that [(a student of mine: y) [y was called before the dean]]]

(b) (each teacher: x) [(a student of mine: y) [x thinks that [y was called before the dean]]]

(c) (a student of mine: y) [(each teacher: x) [x thinks that [y was called before the dean]]]

The three degrees of the referential scale we offered in **4.1** match the three readings: (34a) is non-specific, (34b) is vaguely specific and (34c) is strongly specific on the basis of our definitions. It seems that this referential perspective can account for ambiguities arising from different scopes of quantifiers as well. On the other hand, our discussion makes it clear that specificity cannot be explained only by establishing different scope relations.

Summarizing our results, we saw that though specificity can be reflected in the syntax, its many-sided character cannot be accounted for on syntactic grounds. It does not belong solely to pragmatics either, although a semantic framework flexible enough to characterize specificity is able to incorporate the relevant aspects of pragmatic approaches.

References

Diesing, M. 1992. Indefinites. MIT Press, Cambridge, MA.

Enç, M. 1991. The semantics of specificity. In: Linguistic Inquiry 22: 1–25.

Fodor, J.D.–Sag, I. 1982. Referential and quantificational indefinites. In: Linguistics and Philosophy 5: 355–98.

Groenendijk, J.–Stokhof, M. 1980. A pragmatic analysis of specificity. In: Heny, F. (ed.): Ambiguities in Intensional Contexts, 153–90. Reidel Publishing Company, Dordrecht.

Landman, F. 1986. Pegs and Alecs. In: Towards a Theory of Information: The Status of Partial Objects in Semantics. Foris Publications, Dordrecht.

Ludlow, P.–Neale, S. 1991. Indefinite descriptions: In defense of Russell. In: Linguistics and Philosophy 14: 171–202.

Reuland, E. 1988. Indefinite subjects. In: Proceedings of NELS 18: 375–94. GLSA, University of Massachusetts, Amherst.

Addresses of the authors: Péter Ábel
Department of General Linguistics
Attila József University
Egyetem u. 2.
H–6701 Szeged
Hungary
e-mail: h031209@huszeg11.bitnet

Márta Maleczki
Department of General Linguistics
Attila József University
Egyetem u. 2.
H–6701 Szeged
Hungary
e-mail: h3738mal@ella.hu

Acta Linguistica Hungarica, Vol. 42 (3-4), pp. 227-245 (1994)

OUTLINE OF AN INFORMATION-FLOW MODEL OF GENERICS*

LAWRENCE CAVEDON-SHEILA GLASBEY

1. Introduction

We begin with a brief discussion of the existence of multiple readings for some generic sentences, showing how Krifka (1995) represents the semantics of such sentences, using the quantificational account of generics developed by Wilkinson (1986, 1991), Diesing (1988) and Kratzer (1995) and a modal account of the semantics of the generic operator. Krifka's account, we will show, does not deal satisfactorily with the observation that the truth of a generic sentence may be relative to a particular context. For example, some pairs of mutually conflicting generics are simultaneously acceptable, even by the same speaker. In such cases, the potential conflict can be resolved by viewing each generic as asserting a claim involving a particular context.

We will show how this problem may be addressed using a model of generics based on Barwise and Seligman's **Channel Theory**, a recently-developed mathematical model of information-flow (Barwise 1993; Barwise–Seligman 1994; Seligman–Barwise 1993). Channel Theory gives a structural account of law-like regularities, in particular, encompassing a notion of context and allowing exceptions for any regularity. The Channel-Theoretic account involves several other properties that are desirable to an analysis of generics, and these are also discussed. In particular, a logic for reasoning with generics is described. This logic is obtained directly from Cavedon's (1994) Channel-Theoretic logic of conditionals and supports many important patterns of reasoning. We close by providing a detailed analysis of Krifka's examples within the Channel-Theoretic framework, using Situation-Theoretic tools to model the types involved.

* The authors would like to thank Robin Cooper, Jerry Seligman and participants of the Fifth Symposium of Logic and Language for many useful comments. Sheila Glasbey's work is supported by an EPSRC postdoctoral fellowship.

2. Generic sentences

Krifka (1995) draws attention to the existence of at least two readings for the generic sentence:[1]

(1) *Mary smokes after dinner.*

One reading can, according to Krifka, be roughly characterised as "In after-dinner situations involving Mary, she usually smokes". We will call this the (a) reading. The (a) reading can be seen as an answer to the question "What does Mary do after dinner?" A second reading he characterises as "When Mary smokes, it is usually in after-dinner situations", and can be seen as an answer to the question "When does Mary smoke?". We will call this the (b) reading.[2]

It was the existence of more than one reading for some generic sentences that motivated in part the proposal that a dyadic operator is needed to capture the semantics of generics (see Krifka *et al.* 1995 and Carlson 1989) for further explanation). Carlson (1977) originally proposed a one-place generic operator, which applied to the predicate part of the sentence only. While this accounted for a great deal of data, it failed to predict the existence of multiple generic readings for sentences such as:

(2) *Hurricanes arise in this part of the Pacific.*

which has one reading which says that hurricanes in general arise in this part of the Pacific (which of course is unlikely given our knowledge of hurricanes) and a second reading which says that it is a general property of this part of the Pacific that hurricanes arise there.

Wilkinson (1986, 1991) proposes in place of Carlson's (1977) analysis what she calls a **quantificational** account of generic sentences involving indefinites. The generic operator, which she calls GEN, is viewed as an unselective quantifier in the manner that Lewis (1975) treats adverbs like 'usually'. GEN can thus bind any variables in its scope. Any variables that are not bound by GEN are bound by existential closure, in the manner of Heim (1982). GEN relates two elements of the sentence, the **restrictor** and the **matrix**. Variables in the restrictor are bound by GEN, and variables that appear only in the matrix are bound by existential closure.

[1] This type of generic sentence, with a single subject, is often known as a habitual.

[2] Krifka points out, too, that intonational features such as stress placement play a part in distinguishing between the readings.

Using the above framework, Krifka represents the (a) and (b) readings of (1) as below. We give here his preliminary definition, which he refines later in his paper to accommodate a dynamic framework of interpretation based on Rooth (1987) and the 'alternative semantics' of Rooth (1985). We give the earlier version for clarity, as it does not affect the point we wish to make.

Reading (a): $\mathrm{GEN}[x,s;](x = \mathrm{Mary}\ \&\ \text{after-dinner}(s)\ \&\ \mathrm{in}(x,s);\ \mathrm{smoke}(x,s))$

The convention is that those variables to the left of ';' in the first argument are interpreted as bound by GEN, and any variables to the right of ';' are interpreted as existentially quantified—thus, x and s are both bound by GEN here. We may choose to interpret this formalism as meaning that the (a) reading is somehow making a general statement about after-dinner situations involving Mary. This, of course, ties in with viewing this reading as an answer to a question about what Mary does after dinner. Thus, in some sense, "after dinner situations involving Mary" may be seen as the topic of the sentence on this reading. Krifka identifies the restrictor/matrix partition employed by Wilkinson *et al.* with the background/focus distinction which has been widely discussed in the literature. While this identification may appear to be an intuitive one, it is not generally accepted (nor is there one single clear notion in the literature of exactly what the background/focus distinction is and whether or not it should be viewed as recursive). Further work is required, but we will not pursue this here. (See Partee (1991) for extensive discussion of these and related issues.) Rather, we will regard the generic operator as interacting with the restrictor/matrix partition, without taking Krifka's step of identifying this with background/focus.

Reading (b): $\mathrm{GEN}[x,s;](x = \mathrm{Mary}\ \&\ \mathrm{smoking}(x,s);\ \text{after-dinner}(s))$

We saw above that reading (b) is characterised intuitively as "When Mary smokes, it is usually in after-dinner situations". Krifka formalises this "usually" by providing a semantics for GEN using what could be called a **modal logic of normality.**[3] This involves an accessibility relation $\leq_w$ (defined with respect to a world w^4) between worlds which is interpreted as a normality ordering:

[3] Related logics have recently been used by such authors as Boutilier (1992) and Morreau (1992) to provide logical accounts of generics and default reasoning.

[4] Actually, the accessibility relation is also relative to a **modal background**, but this concept is ignored here for the sake of simplicity and brevity.

$v \leq_w u$ if and only if v is "as normal as" u. The precise details of the truth-conditions of GEN are omitted here but, informally, GEN$(A; B)$ holds at w if for every world u at which A holds, there is a world v as normal as u, such that for every world v' satisfying A (where v' is as normal as v), v' also satisfies B. That is, as one approaches a limit of "maximal normality", all increasingly normal worlds that satisfy A also satisfy B.

Logics of normality possess many features that prove desirable in a semantics for generics.[5] However, this approach has a problem with the (b) reading of the sentence above. Observe that it is possible to follow (1) on reading (b) with (3):

(3) *She smokes before breakfast too.*

Adding (3) in no way appears to cancel or override the information given in (1). The impression is simply one of further information being supplied, in order to give a more complete specification of those times of the day when Mary habitually smokes. (1) and (3) could be uttered sequentially by the same speaker, for example, in answer to the question "When does Mary smoke?".

Indeed, it would be possible for the speaker to go on to add (4):

(4) *And she smokes after lunch.*

and so on. The speaker is simply adding more information and working towards a more complete answer to the question.

It may be questioned whether it is still reading (b) that is involved in this case, or whether it is in fact reading (a) that we have here. But if it were reading (a), then the speaker would be conveying that in the maximally normal "after-dinner" situations, Mary smokes. This does not seem to be correct. If a speaker utters (1) in reply to the question "When does Mary smoke", then he is not claiming that Mary smokes in the most normal after-dinner situations, but rather he is making what is in one sense a far weaker claim that may be true even if in the majority of after dinner situations Mary does not smoke. So it seems that reading (b) is still the relevant one here.

Another possibility is that (1, 3 and 4) are uttered by different speakers who have knowledge of Mary's behaviour at different times of day. For example, her husband might only be aware of Mary's smoking before breakfast, and her

[5] Morreau (1992) discusses these in detail.

children, with whom she generally eats dinner, might only be aware that she smokes after dinner.

Krifka's analysis of (1) on reading (b) is no longer correct once (3) has been uttered. It need no longer be true that when Mary smokes it is usually (or "in the most normal cases") "after dinner". It might just as well be "before breakfast". Yet there is no sense of the speaker correcting himself, or of the second speaker correcting the first. There is, rather, a sense of disjunction—a hearer who has heard (only) (1) and (3) may conclude, if he sees Mary smoking, that it is probably *either* after dinner or before breakfast, if he has no information to rule out one or other of these disjuncts. It should be pointed out that the fact that the normality ordering is relative to a world and modal context does not alleviate the problem. We would expect the (b) readings of both (1) and (3) (i.e. with the consequent being either *after-dinner(s)* or *before-breakfast(s)*) to be supported in the same world, and with respect to the same modal context. Thus it appears that Krifka's analysis must be modified.

What we wish to be able to say is that reading (b) of (1) conveys that situations of a kind where Mary is smoking[6] are somehow connected to situations that are temporally located "after dinner" **in a particular context**. But this does not prevent situations of the kind where Mary is smoking from being connected to situations of a kind located at other times of day, such as "before breakfast", in another context or viewed from a different perspective. It is not clear how to capture this context sensitivity or "perspectivity" in the possible worlds framework of normality logics.

In the next section, we outline **Channel Theory**, a mathematical model of information-flow currently being developed by Seligman and Barwise (Barwise 1993; Barwise–Seligman 1994; Seligman–Barwise 1993), which we then use as a basis for a formal model of generics which allows us to deal directly with the problem of context.

3. Channel Theory

Channel Theory is an account of **regularities** (i.e. law-like dependencies) that builds on the Situation Theoretic concept of a **constraint**.[7] Barwise and Seligman adopt a realist stance towards regularities, taking them to be part of the

[6] We use the progressive merely to indicate that the situation corresponds to a single event or "episode" of Mary smoking, not a generalisation over such episodes.

[7] Ter Meulen (1986) has presented a model of generics based on Situation Theoretic **constraints**, a framework which can be seen as a precursor to Channel Theory.

natural order of the world. As such, Channel Theory does not provide a reductionist account of regularities—they are taken to be primitive intensional objects of the theory. However, Channel Theory does provide a **structural** account of regularity—a regularity is decomposed into a type-level **constraint**, and **connections** between tokens relevant to those types. This two-level structure is used to account for the way in which a regularity can have exceptions without invalidating the general rule. This property is exactly the sort that is crucial to an adequate formal model of generic sentences.

In our model of generics, the content of a generic is directly associated with a regularity. However, any such regularity includes as a component a **channel**—channels are objects that support regularities (in the sense of making them "true") under the two-level analysis outlined above. In Barwise and Seligman's framework, a regularity can only be true or false[8] **relative to some channel**. Any channel that supports a regularity then plays the role of a **context** supporting that regularity, which leads to an analysis of generics that solves the problem described in the previous section.

At this point, we define the main formal concepts of Channel Theory. At the heart of Channel Theory is a relativistic theory of information, based on the concept of a **classification**. A classification is a structure that carves up (part of) the world into **tokens** and assigns various **types** to these tokens.

Definition: A **classification** A is a structure $\langle tok(A), typ(A), :_A\rangle$ consisting of a set of **tokens** $tok(A)$, a set of **types** $typ(A)$ and a relation $:_A$ on $tok(A) \times typ(A)$. For $t \in tok(A)$ and $\phi \in typ(A)$, we say t is **classified by** ϕ **in** A if $t :_A \phi$ holds.[9]

The tokens of a classification may be objects, individuals, situations or even other classifications, while the types are properties appropriate for classifying (along some dimension) these tokens. A classification can be seen as a representation of information of a particular sort. This representation is highly relativistic—intuitively, one could view a classification as being relativised to a specific agent, epistemic state, point of view, or perspective on the world.

The second main concept is that of a **channel**. A channel C is an informational link between two classifications A and B, licensing the flow of infor-

[8] It is a bit of an abuse to talk of truth when dealing with Channel Theory since it is concerned with **information** and **information-flow**, rather than truth. However, it simplifies the presentation to do so.

[9] In general, there are two classification relations, one for positive and one for negative classification. However, we do not make use of negative classification in this abstract, so it is omitted for simplicity. We drop the subscript from the classification relation when this causes no confusion.

mation of a particular sort. Formally, a channel is itself a classification—the tokens of C are **connections**, each linking tokens of A and B.[10]

Definition: Let A and B be classifications. A channel $C : A \Rightarrow B$ linking A and B is a classification $\langle tok(C), typ(C), :\rangle$. The tokens of C are **connections**, denoted $s \mapsto s'$, with $s \in tok(A)$ and $s' \in tok(B)$. The types of C are *constraints*, denoted $\phi \rightarrow \psi$, with $\phi \in typ(A)$ and $\psi \in typ(B)$.

The basic intuition of this definition is the following. $(t \mapsto t' : \phi \rightarrow \psi)$ holding in C signifies that t being of type ϕ carries the information that t' is of type ψ. The type-token distinction allows for a characterisation of **exception** to the general rule—a type-level constraint can be supported in a channel even though there are connections in that channel which are not classified by it. As a simple example, consider an electrical circuit involving a (constant) power source, a switch and a light bulb. A classification A involving the switch consists of the token s and types *open* and *closed*. A classification B involving the bulb consists of the token b and types *on* and *off*. An obvious channel is $C : A \Rightarrow B$, containing the connection $s \mapsto b$ as a token and the constraint *closed* $\rightarrow$ *on* as a type. Of course, there may be other constraints and other channels. If, under some abnormal condition, the circuit fails to function as expected, then the connection between the switch and the bulb has failed—this is modelled by the connection $s \mapsto b$ being an exception to *closed* $\rightarrow$ *on*.

It should be noted that, within Channel Theory, there is no direct relationship imposed between the connections of a given channel and the assertibility of its constraints—as far as the theory goes, every connection may well constitute an exception, without invalidating the constraints within that particular channel. This is, of course, reminiscent of the relation between a generic property and its instances.[11]

A simple Channel Theoretic model of generics[12] can be defined by using the type-level of a classification to model properties and the token-level to model individuals. A type-level constraint is used to model a generic relationship, while a connection between a token t and itself is included in a channel if it is appropriate to the generic relationship represented by some constraint in

[10] We slightly abuse Seligman–Barwise's definition here—connections and constraints should not be identified with the tokens/types they connect. However, this abuse simplifies matters somewhat.

[11] For example, Morreau (1992) points out that "Potatoes contain vitamin C" is true even if we take all the potatoes in the world and boil them until they no longer do so.

[12] We are solely concerned here with what Krifka *et al.* (1995) call **characterising sentences**.

that channel. For example, the sentence "Swans are white" describes a channel C containing connections of the form $s \mapsto s$ and a constraint $swan \rightarrow white$, where s is taken from some set of swans, $swan$ is a type that holds of swan-objects and $white$ is a type that holds of white-objects. Of course, it may be the case that a particular connection $s \mapsto s$ is not classified by this constraint in C: if $(s : swan)$ holds, then this represents the case whereby s is an exception to the generic relation. Note that support for a given generic is only ever with respect to some channel. As such, the channel can be seen as the **demonstrative content** of the generic.[13]

4. Some general properties of the Channel Theoretic model

In this section, we briefly discuss some of the general properties of the Channel Theoretic model of generics. These issues are discussed in greater detail in Cavedon (1995).

4.1. Generics, channels and regularities

The analysis of generics sketched above inherits several desirable properties from Channel Theory's two-level representation of regularities. The fact that the truth of a type-level regularity (with respect to a channel) does not depend on the properties of the associated individuals avoids several well-known problems. For example, it could be argued that "John shoots burglars" is acceptable even if John keeps a gun by the bed and intends to use it but has never had occasion to do so. In the Channel Theoretic analysis, the fact that there has never been a **burglar** situation does not preclude there being a regularity between situation-types where John encounters a burglar and situation-types where John shoots any such intruder. The separation of regularities into type and token levels allows for exceptions to a general regularity to be handled without compromising the truth of that regularity. Of course, if one wants to give an account of **why** a particular regularity is supported, then Channel Theory does not provide anything more informative than the claim that that is the way the world is. This is sufficient for our concerns here—in any case, the Channel Theoretic analysis can be seen as consistent with any particular

[13] For a more satisfactory account of generics, we really need a much more complex representation of tokens and types. We show below how the complex types are constructed. Theoretically, a semantics for the general case of the GEN operator could be provided using the Channel Theoretic model. However, we do not have the space here to present all the necessary formal machinery. This issue is discussed in Cavedon (1995).

reductionist account of regularity one may want to use (e.g. a normative one), without requiring a commitment to any such account.[14]

The issue we are more concerned with in this paper is the **context** that is provided for each generic by the Channel Theoretic analysis. A generic is only supported with respect to a channel, i.e. if the corresponding type-level regularity is a type of that channel. Any channel C is itself a classification, with a set of tokens consisting of connections between individuals relevant to the type-level regularities. Individuals that are not relevant to the type-level regularities supported by C do not participate in connections at the token-level of C. In fact, such individuals may not even be tokens in the classifications linked by C—i.e. such individuals are outside the scope of applicability of the regularities in C. This illustrates how C provides a context for a generic that describes a type-level regularity in C—the applicability of the generic determines the set of tokens with which C is concerned, and any claim made by the generic is limited to this set of tokens.

For example, consider the following pair of sentences:

(5) (a) Peacocks lay eggs.
(b) Peacocks have brightly coloured tail-feathers.

Such examples are problematic for logics of normality—since only female peacocks lay eggs and only male peacocks have brightly coloured tail-feathers, there can be no "normal" peacock individuals. Morreau (1992) claims that this problem is simply an artefact of the translation into an appropriate logical form—i.e. such a form should contain *male* or *female* in the restrictor—but the Channel Theoretic analysis seems to provide a neater solution: each of these generics is supported by a different channel, reflecting different contexts. The context of (5a) is concerned solely with female peacocks while that of (5b) is concerned solely with male peacocks. Of course, these would be the contexts associated with the content of the assertion of the speaker who utters these generics. A hearer may ascribe a different context to the generics in her interpretation (possibly leading to a perceived inconsistency). Such interesting possibilities can be modelled in a Channel Theoretic analysis of communication, but this is outside the scope of the current paper.[15] The treatment of the

[14] One potential problem for accounts of generics that allow such an independence between type-level and token-level is that it does not easily lead to a logic of **default reasoning**—i.e. a logic for (defeasibly) reasoning about the properties of individuals. This issue is discussed later.

[15] See Healey–Vogel (1994) for a Channel Theoretic model of dialogue.

sentences of section **2** proceeds along similar lines to the above—we return to these sentences in section **5**.

It should be noted that the contextual element in the semantics of a generic does not preclude an analysis of **nested** generics (i.e. generics for which the premise or consequent is itself a generic). A channel links two classifications. However, since any channel is itself a classification, a given channel C, where C is the demonstrative content of some generic, may link a classification to some other channel C', where C' is itself the demonstrative content of some other generic. This allows an analysis of nested generics to be handled without any extension to the basic framework.

4.2. Reasoning with generics

An important component of a formal model of generics is the ability to reason with them, allowing both the inference of new generic relationships and the default inference of properties of individuals. Seligman and Barwise (1993) define a number of operations on channels which allow new channels to be defined from old. For example, a **serial composition** operation (.;.) combines two channels in a transitive way—if C supports the regularity "Birds fly" and C' supports "Sparrows are birds", then $(C';C)$ supports "Sparrows fly". The logic that results from Seligman and Barwise's operations is not suitable as a logic of generics—for example, it supports Monotonicity, Transitivity and Contraposition, patterns of inference that are invalid for generics (e.g. see Morreau (1992)). Cavedon (1994) refines the channel operations so that they account for **background conditions**—these are the "normal conditions" that are assumed to hold in a channel that supports a given regularity. Crucially, Cavedon's framework does not require channel's background condition to be represented in that channel—instead, they are captured by the way in which the channel is related to other channels in a **channel hierarchy**. Using this framework, Cavedon obtains a powerful logic of conditionals that invalidates the unwanted patterns of inference mentioned above.

Using Cavedon's framework, we have obtained a logic for generics directly from the logic of conditionals. This is related to work by Boutilier (1992) and Morreau (1992) who each obtain logics of normality by using possible-worlds conditional logics as a starting point.[16] The Channel Theoretic logic, however, requires no extra machinery to the logic of conditionals itself—a simple closure assumption regarding the associated channel hierarchy is sufficient to ensure

[16] Krifka's (1995) logic can be seen in the same light, using a different conditional logic as its basis.

the required behaviour. This logic has many desirable properties. In particular, it satisfies Morreau's (1992) **Penguin Principle**—from the generics "Birds fly", "Penguins are birds" and "Penguins can't fly", one is led unambiguously to the conclusion that penguins are unable to fly (the "more specific" information overrides the transitive inference).[17]

A logic for default reasoning involves not only a logic for the generic relationships themselves, but also a methodology for (defeasibly) reasoning about the properties of individuals themselves. To be able to draw any conclusion regarding a token-level individual t, it is of course necessary to assume that t falls within the domain of the generic relationship. In Channel Theoretic terms, this involves assuming that the connection involving t is classified by the constraint representing the generic in question, within the appropriate channel. By defining a suitable **maximal normality condition**, which basically imposes the assumption that there are as few exceptions as possible within a channel, a logic for reasoning about individuals is obtained. This logic satisfies all the Asher–Morreau (1991) principles of generic and default reasoning.

Details of the logics of generics and default reasoning can be found in Cavedon (1995).

5. Application to Krifka's example

We now return to Krifka's problematic example and show how a Channel Theoretic analysis addresses the issues we raised.

Each channel can be intuitively seen as corresponding to a particular "context" or "perspective". This gives us a way to express reading (b) of (1). In one channel (call this C_1), a situation-type where Mary is smoking at a time T on a day D indicates that T is temporally included in the *after-dinner* period of day D. We can express this informally as a constraint in $typ(C_1)$ as follows:

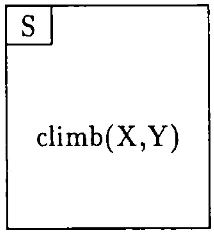
$$smoking(mary, T, D) \rightarrow T \sqsubseteq \textit{after-dinner}(D),$$

where $\sqsubseteq$ conveys temporal inclusion.

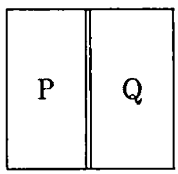
In order to express formally the above constraint, we use the notation EKN (extended Kamp notation) introduced in Barwise–Cooper (1993), which

[17] Morreau (1992) himself has problems with the Penguin Principle when all three are generics are defeasible (in the example given, "Penguins are birds" is *not* defeasible). The Channel Theoretic logic has no such problems.

is a graphical notation for Situation Theory. A box notation is used for Situation Theoretic objects such as infons, situations and propositions, based upon the graphical notation of DRT. EKN boxes look rather like DRSs, but one important difference is that EKN boxes may contain situations. Propositions in EKN include objects of the form:



which is the proposition that a situation S supports an infon *climb*(X, Y).[18] Situation Theoretic objects may have restrictions imposed on them, as shown below, where the object P is restricted by a proposition or conjunction of propositions Q:



P denotes an object only if Q is true. A **type** is formed by simultaneously abstracting over the parameters of the proposition, using the technique of **simultaneous abstraction** developed in Aczel–Lunnon (1991) and Lunnon (1991). The roles corresponding to the abstracted parameters can be indexed by any object, and are indexed here by r_1 to r_n.

Using this notation we can express the constraint corresponding to reading (b) as in Fig. 1:

[18] S, X and Y are parameters, denoted by capital letters.

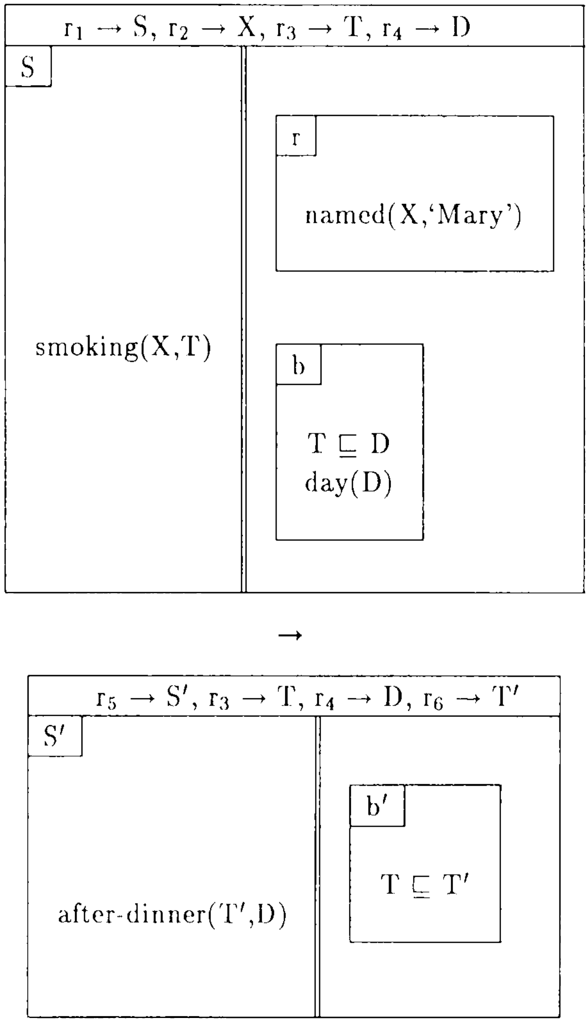


Fig. 1
Constraint corresponding to reading (b) of (1)

Here, *r* is the resource situation supporting information about the naming of Mary, and *b* is a background situation supporting information about the time relations. (For further explanation of resource and background situations, see Barwise–Cooper (1993).)

The above tells us that there is a constraint in channel C_1 between two types. The antecedent type is a 4-ary type of a situation S, an individual X named Mary, a time T and a day D, where S supports the fact that Mary smokes at T. The succedent type is a 4-ary type of a situation S', the same time T, the same day D and a time T' that temporally includes T, such that S' supports the fact that T' is the (whole) after-dinner period of day D. A similar constraint (in C_2) could be expressed which tells us that this same succedent type indicates a type where T' is the before-breakfast period of day D.

We can consider (3) as adding the information that there is a second channel (C_2) where situations of a type where Mary is smoking at T on day D indicate that T is included in the *before-breakfast* period of day D—i.e. C_2 contains the constraint expressed informally as:

$$smoking(mary, T, D) \rightarrow T \sqsubseteq \textit{before-breakfast}(D).$$

The fact that we can introduce as many new channels as we like explains the observation that the speaker (or another speaker) may go on adding further times of day during which Mary habitually smokes. Note that the existence of a constraint between two situation-types in no way forces "most" situations satisfying the antecedent type to be connected to situations satisfying the succedent type—thus, we overcome the problem inherent in Krifka's analysis, whereby (2) appears to contradict (1). Interestingly, given channels C_1 and C_2, the Channel Theoretic logic described in section **4.2** allows us to infer the existence of a composite channel C' that supports the generic represented by the following constraint (expressed informally):

$$smoking(mary, T, D) \rightarrow (T \sqsubseteq \textit{after-dinner}(D) \vee T \sqsubseteq \textit{before-breakfast}(D))$$

The logic also allows us to arrive at the default conclusion that, given that Mary is smoking, she is either in an *after-dinner* or *before-breakfast* situation.[19]

Reading (a) of (1) can be captured by saying that in some channel C_3 there is a constraint between a situation type where the situation S supports the fact that Mary is in S for a time T, where T is the after-dinner period of day D, and the situation-type where some situation S' supports the fact that Mary is smoking at some time T' that is temporally included within T.

We can express reading (a) formally as in Fig. 2:

[19] The manner in which these conclusions are arrived at is detailed in Cavedon (1992).

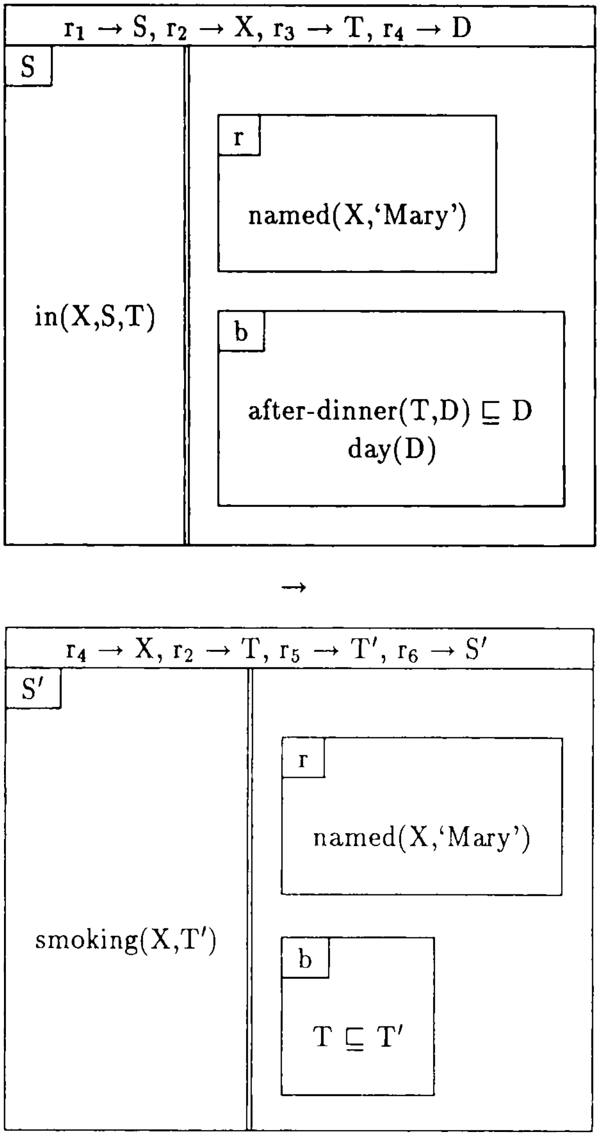


Fig. 2
Constraint corresponding to reading (a) of (1)

This tells us that there is a constraint in a channel C_3 between two types. The antecedent is a 4-ary type of a situation S, an individual X named Mary, a time T which is the after-dinner period of D, and a day D, where S supports the information that X is in that situation for the duration of time T. The succedent is a 4-ary type of the individual named Mary, the same time T, a time T' temporally-included in T and a situation S'.

By expressing the constraint in this way we avoid a problem for Krifka's account which concerns the "size" (i.e. temporal duration) of the situations in which we are interested. His representation of reading (a) effectively says that in the most normal after-dinner situations involving Mary, she smokes. But if a situation can be of any temporal duration whatsoever, then there will be many situations of small temporal duration in which Mary does not smoke—but what is there to rule out such situations from being counted among the most normal situations? Yet (1) may still be true on reading (a). It appears that if we are to use situations in this way we must somehow restrict their size.[20] Our above formalisation of the constraint does not suffer from this problem, however, because it is only situations which have a temporal duration of the whole of the after-dinner period T which signal situations where Mary smokes. (For a fuller discussion of how the temporal duration of situations is related to the time arguments of the infons they support, see Glasbey (1994, Ch. 3).)

6. Conclusions and further work

We believe that Channel Theory provides a very sound basis for a formal model of generics. The contextual component of a channel allows an explanation of aspects of generics that have proved problematic in the past. The separation of regularities into type-level and token-level components means that the truth of a generic is independent of the behaviour of its instances, allowing an account of exceptions to the general rules to be easily given. Perhaps most interestingly, Channel Theory allows us to take a completely neutral, non-reductionist stance on how generics arise, while still allowing the model to be employed to make useful predictions, regarding both the derivation of new generic relationships and the default inference of information regarding instances of these generics. We believe that this lack of commitment without the sacrifice of reasoning power is extremely attractive.

[20] It should be pointed out, too, that introducing the idea, as Krifka does later in his paper, that "in the most normal situations involving Mary *if she does some alternative to smoking* (where the set of alternatives is contextually defined) then she smokes" does not address this problem.

There is much further work that we intend to pursue. Section **5** presented a detailed analysis of both readings of example (1), using tools from Situation Theory. This allows the use of a very expressive semantic theory in the analysis of the restrictor and matrix of a generic. It would be particularly interesting to explore how the stage-level/individual-level distinction of Carlson (1977), which Kratzer (1995) proposes may correspond to the presence/absence of a location or event argument to the predicate, may be represented in the Situation Theoretic formalism we employ here. We would need to investigate whether or not we can use situations to correspond to Kratzer's 'events', or whether perhaps we require to make a distinction between situations which are spatiotemporally located and situations which are not (see Glasbey (1994b) for discussion). Further work is also required to extend the Channel Theoretic semantics to the full expressive notation allowed in the use of the GEN operator. The current semantics does not properly cater for variables that are existentially quantified within the matrix of the operator. Another question, mentioned earlier, regards the precise relationship between the restrictor/matrix partition of a generic and the background/focus division (see Partee (1991) for discussion of some relevant issues). We intend to investigate, too, how the semantics of generic sentences like "Mary smokes", where contextual restrictions clearly must be built into the representation in some way, might be addressed using the Channel Theoretic model of generics presented here. Finally, a full investigation of the formal properties of the logic of generics and default reasoning described in section **4.2** is a further important topic.

References

Aczel, P.–Lunnon, R. 1991. Universes and parameters. In: Barwise, J.–Gawron,–J.M. Plotkin, G.–Tutiya, S. (eds): Situation Theory and its Applications. Vol. 2. CSLI, Stanford.

Asher, N.–Morreau, M. 1991. Commonsense entailment: a modal theory of commonsense reasoning. In: Proceedings of the International Joint Conference on Artificial Intelligence. Sydney, Australia.

Barwise, J. 1993. Constraints, channels and the flow of information. In: Aczel, P.–Israel, D.–Katagiri, Y.–Peters, S. (eds): Situation Theory and its Applications. Vol. 3. CSLI, Stanford.

Barwise, J.–Cooper, R. 1993. Extended Kamp notation: a graphical notation for Situation Theory. In: Aczel, P.–Israel, D.–Katagiri, Y.–Peters, S. (eds): Situation Theory and its Applications. Vol. 3. CSLI, Stanford.

Barwise, J.–Seligman, J. 1994. The rights and wrongs of natural regularity. In: Tomberlin, J. (ed.): Philosophical Perspectives. Vol. 8. Ridgeview, California.

Boutilier, C. 1992. Conditional logics for default reasoning and belief revision. Ph.D. thesis. Department of Computer Science, University of Toronto.

Carlson, G. 1977. Reference to kinds in English. Ph.D. thesis. University of Massachusetts.

Carlson, G. 1989. On the semantic composition of English generic sentences. In: Chierchia, G.-Partee, B.-Turner, R. (eds): Properties, Types and Meaning. Vol. 2: 167-92. Kluwer, Dordrecht, Holland.

Cavedon, L. 1994. A Channel-Theoretic model for conditional logics (extended abstract). In: Conference on Information-Oriented Approaches to Logic, Language and Computation. Moraga, CA.

Cavedon, L. 1995. A Channel-Theoretic approach to conditional reasoning. Ph.D. thesis. Centre for Cognitive Science, University of Edinburgh.

Diesing, M. 1988. Bare plural subjects and the stage/individual contrast. In: Krifka, M. (ed.): Genericity in Natural Language. Proceedings of the 1988 Tübingen Conference. Seminar für Natürlich-Sprachliche Systeme der Universität Tübingen.

Glasbey, S. 1994a. Event structure in natural language discourse. Ph.D. thesis. Centre for Cognitive Science, University of Edinburgh.

Glasbey, S. 1994b. Progressives, events and states. In: Dekker, P.-Stokhof, M. (eds): Proceedings of the Ninth Amsterdam Colloquium. Vol. 2: 313-32. Department of Philosophy, University of Amsterdam, Institute for Logic, Language and Computation.

Healey, P.-Vogel, C. 1994. A Situation Theoretic model of dialogue. In: Jokinen, K. (ed.): Pragmatics in Dialogue Management. Gothenburg Monographs in Linguistics.

Heim, I. 1982. The semantics of definite and indefinite noun phrases. Ph.D. thesis. University of Massachusetts, Amherst.

Kratzer, A. 1995. Stage-level and individual-level predicates. In: Pelletier, F.-Carlson, G. (eds): The Generic Book. University of Chicago Press.

Krifka, M. 1995. Focus and the interpretation of generic sentences. In: Pelletier, F.-Carlson, G. (eds): The Generic Book. University of Chicago Press.

Krifka, M.-Pelletier, J.-Carlson, G.-ter Meulen, A.-Link, G.-Chierchia, G. 1995. Introduction. In: Pelletier, F.-Carlson, G. (eds): The Generic Book. University of Chicago Press.

Lewis, D. 1975. Adverbs of quantification. In: Keenan, E. (ed.): Formal Semantics of Natural Language. Cambridge University Press.

Lunnon, R. 1991. Generalized universes. Ph.D. thesis. University of Manchester.

Morreau, M. 1992. Conditionals in philosophy and artificial intelligence. Ph.D. thesis. Department of Philosophy, University of Amsterdam.

Partee, B. H. 1991. Topic, focus and quantification. In: Moore, S.-Wyner, A. (eds): Semantics and Linguistic Theory I. Cornell Working Papers in Linguistics.

Rooth, M. 1985. Association with focus. Ph.D. thesis. University of Massachusetts at Amherst.

Rooth, M. 1987. Noun phrase interpretation in Montague grammar, file change semantics, and situation semantics. In: Gärdenfors, P. (ed.): Generalized Quantifiers. Reidel, Dordrecht.

Seligman, J.-Barwise, J. 1993. Channel Theory: towards a mathematics of imperfect information flow (draft). Available by ftp as phil.indiana.edu:/pub/SelBar93.ps.

ter Meulen, A. 1986. Generic information, conditional contexts and constraints. In: Traugott, E.C. – ter Meulen, A. – Reilly, J.S. – Ferguson, C.A. (eds): On Conditionals. Cambridge University Press.

Wilkinson, K. 1986. Genericity and indefinite NPs. University of Massachusetts at Amherst.

Wilkinson, K. 1991. Studies in the semantics of generic noun phrases. Ph.D. thesis. University of Massachusetts at Amherst.

Addresses of the authors: Lawrence Cavedon
Department of Computer Science
Royal Melbourne Institute of Technology
GPO Box 2476V, Melbourne 3001, Australia
e-mail: cavedon@cs.rmit.edu.au

Sheila Glasbey
Centre for Cognitive Science
University of Edinburgh
2 Buccleuch Place, Edinburgh EH8 9LW, UK
e-mail: srg@cogsci.ed.ac.uk

Acta Linguistica Hungarica, Vol. 42 (3-4), pp. 247-252 (1994)

A COMPOSITIONAL SEMANTICS FOR THE SPATIOTEMPORAL PROPERTIES OF MOTION VERBS AND SPATIAL PPS IN FRENCH

NICHOLAS ASHER–PIERRE SABLAYROLLES

This study formalizes the spatial and temporal information that is contained within the meaning of verbs of movement and sentences constructed from them. It falls within the scope of the study of the model theoretic semantics of natural languages. Many studies have already been made concerning the problems of representation of spatial and temporal knowledge and of reasoning about these knowledges in Artificial Intelligence, especially in the field of robotics. But, in view of the complexity of the representation problem, there has been for some years a demand for more cognitive theories of representation and reasoning, i.e. theories based on the way human beings represent in their mind this knowledge and perform highly complex reasoning on the representations they build. A methodology to understand how human beings process spatiotemporal information is to see how this information is reflected in natural language. This is our aim here. We are principally interested in spatial and temporal information, so, we will focus on those expressions in natural language which describe motion. Motion, representing a succession of spatial locations in time, offers a rich field of interrelated concepts having to do with spatial and temporal information.

In this paper, we define a compositional semantics for the spatiotemporal properties of motion complexes in French, which we define here as a motion verb followed by a spatial PP, and on the basis of this semantics we offer a typology of motion verbs and spatial prepositions. We specify the basic semantics of motion verbs and prepositions in terms of their relations to three types of spatiotemporal entities: locations, positions, and postures. Locations are functional portions of space that are lexicalizable; positions are portions of a surface inside a location, without functionality or associated lexical item; postures comprise relations between the parts of an entity. For us, a motion verb involves a change in either the location, position or posture of one or more objects (though for our study we have restricted ourselves to contexts with one moving "target"). We also assume that each eventuality described by a motion verb has three distinguished spatiotemporal positions—a source,

1216-8076/94/$ 5.00

a strict-inner-path (SIP) and a goal with respect to each of our ontological categories, location, position and posture. For instance, the source and goal locations of an entity are the starting and the ending locations of a motion respectively, while the SIP location is the sequence of locations the moving entity follows along its motion.

These ontological categories allow us to avoid certain difficulties in earlier studies (e.g. in Boons 1985) with the concept of "location", which was not properly defined. For instance, we are able to distinguish the spatiotemporal meanings of the minimal pairs *entrer* 'to go in' / *sortir* 'to go out' on the one hand and *sortir* 'to go out' / *partir* 'to leave' on the other. Further, our ontological assumptions also allow us to compute the right meanings for the motion complexes *sortir de/par le/dans le jardin* 'to go out of/by/into the garden'.

The linguistic study we present here is fully formalized in Asher–Sablayrolles (1994) in a mereologically based theory (Lesniewski 1927–31; Clarke 1981; Aurnague–Vieu 1993) that is compatible with most semantical theories. We have exploited DRT of Kamp (1979) and SDRT of Asher (1993) to model the interaction between discourse structure and lexical meaning. Further, we have made use of the weak conditional operator > of DICE, "Discourse in Commonsense Entailment" of Asher–Morreau (1991), also used in Asher *et al.* (1993), Lascarides–Asher (1993) and Asher (1992), representing defeasible or generic rules (A>B means "if A then normally B"), to encode defaults about lexical entries. We label constituents of discourse representation as having certain properties, and reason about lexical properties within Commonsense Entailment.

To be a little more precise, we define four classes of motion verbs:

- the verbs of change of location (CoL), like *entrer* 'to enter';
- the verbs of change of position (CoPs), like *se déplacer* 'to move around';
- the verbs of inertial change of position (ICoPs), like *courir* 'to run': they can be combined with *sur place* 'in place', in contrast with CoPs verbs;
- and the verbs of change of posture (CoPtu), like *s'asseoir* 'to sit'.

The three first classes are not exclusive; the events denoted by ICoPs complex are subset of those denoted by CoPs complex, which in turn are subset of those denoted by CoL complex. Indeed, one cannot change location without changing position; and one cannot change position without inertially changing position. In contrast, one can change location or position without changing posture. The addition to a motion verb of arguments and adjuncts can only make the resulting complex belonging to the same or to an embedded class than the one of the verb.

We give a more detailed analysis for CoL verbs and CoL complexes. We explain how the source, goal and SIP are organized relative to a location of reference (lref) with respect to which the displacement take sense and which is given by the spatial PP. They generally describe displacements going from the outside to the inside of lref, or vice-versa. Nevertheless, a coarse organization of the space in only two zones, the "interior" and the "exterior" of the lref, quickly proves its insufficiency (see Boons 1985; Laur 1991). We offer a finer organization of the space using in addition of an internal zone (Zih for Zone of inner halo), an external zone of natural contact (Zc for Zone of contact), an outside of proximity (Zoh for Zone of outer halo) and a faraway outside (Zom for Zone of outer most). This complexity in the spatial organization is necessary if one wants to represent and exploit the semantics of motion verbs fully. Our class of CoL verbs can be refined on the basis of which zones the moving entity is inside, at the beginning and at the end of its motion. All the possibilities are not lexicalized in French. In a systematic linguistic study we have found 216 lexical entries for French intransitive CoL verbs that have allowed for the definition of 10 subclasses:

- when the moving entity goes from Zom to Zoh like *S'approcher* 'to approach';
- when the moving entity goes from Zom to Zih like *Arriver* 'to arrive';
- when the moving entity goes from Zoh to Zih like *entrer* 'to enter';
- when the moving entity goes from Zoh to Zc like *Se poser* 'to land';
- when the moving entity goes from Zc to Zoh like *Décoller* 'to take off';
- when the moving entity goes from Zih to Zoh like *Sortir* 'to go out';
- when the moving entity goes from Zih to Zoh like *Dévier* 'to deviate' (this type differs from *Sortir* by the fact that the lref is not a location but an "ideal trajectory");
- when the moving entity goes from Zih to Zom like *Partir* 'to leave';
- when the moving entity goes from Zoh to Zom like *S'éloigner* 'to distance oneself from';
- when the moving entity goes from Zoh to Zoh via Zih, like *Passer* 'to cross'.

Further, CoL verbs are classified as either Initial, Medial or Final. Final verbs focus on the goal, Medial on the SIP and Initial on the source. We have followed the same approach with spatial prepositions. Following Laur (1993), we consider simple prepositions (*dans* 'in') as well as prepositional phrases (*en face de* 'in front of').

We have classified 199 such French prepositions into 16 groups using in addition of our zones two other criteria:

- prepositions can be positional (*dans* 'in') or directional (*dans* 'into');
- directional prepositions can be Initial (*de* 'from'), Medial (*par* 'through') or Final (*à* 'to'), depending if they focus on the source, the SIP or the goal.

On the basis of these classifications, we offer compositional rules, formalized in Asher–Sablayrolles (1994) calculating the spatiotemporal properties of the complex from the ones of the verb and the PP. For lake of space, we do not give the rules here, but provide examples as illustration:

(a) *Jean sort dans le jardin.*
'John goes out into the garden.'

(b) *Jean entre dans le jardin.*
'John goes into the garden.'

(c) *Jean court dans le jardin.*
'John runs in/into the garden.'

These three sentences use the same positional preposition (*dans* 'in') which locates the theme inside Zih('garden').

In (a) we have the Initial verb *sortir* 'to go out'. Our combinational rules say that we have a case of non-congruence: the moving entity goes from Zih(lref) to Zoh(lref), but lref remains non-instantiated (it has to be matched with some location given by the discourse (see Asher–Wada (1989) for anaphora resolution principles) or by the context); in addition, we know that the goal is also part of Zih('garden').

In (b) we have the Final verb *entrer* 'to go in'. We have a case of congruence (lref='garden'): the moving entity goes from Zoh('garden') to Zih('garden'). In (c) we have the Medial verb *courir* 'to run'. Our compositional rules predict two possibilities: congruence (John is already in the garden and stays there during his motion *dans*='in'); and non-congruence (John is not in the garden before his motion, but enters it *dans*='into'). We can derive these meanings compositionally and formally from a syntactic structure and our lexical entries for the verbs and prepositions. Interested readers can consult Asher–Sablayrolles (1994, 10) where we show on examples how to derive DRS using the "bottom-up" DRS construction procedure from Asher (1993), in which the function argument structure of the syntax can be followed.

Finally, we have realized a comparative survey of the literature, distinguishing syntactic approaches (Wunderlich 1991; Maienborn 1992; Guillet 1990) from semantic ones (Hayes 1989; Dervillez-Bastuji 1982; Lamiroy 1987; Boons 1987; Laur 1991). Two conclusions emerge: syntactic approaches do not furnish completely precise criteria and do not cover all the so-called motion verbs. Of course such approaches are not useless; they serve for corrobate semantical criteria built on classifications. The second conclusion is important: all these works (including syntactic approaches) arrive at compatible classifications. This convergence of results for classifications built on so different kinds of criteria (syntactical, semantical, conceptual) gives to them a very strong validity. Our classification is compatible with all the others in the literature and moreover is more detailed.

References

Asher, N.–Morreau, M. 1991. Commonsense entailment: a modal theory of nonmonotonic reasoning. In: Proceedings of IJCAI '91.

Asher, N. 1992. Spatiotemporal structure in texts. In: Akama, S. (ed.): Festschrift for Onishi, I. Oxford University Press.

Asher, N. 1993. Reference to Abstract Objects in Discourse. Kluwer Academic Publishers.

Asher, N.–Aurnague, M.–Bras, M.–Vieu, L. 1993. Space, time and discourse. In: Proceedings of the Dagstuhl Seminar on Universals in the Lexicon; also in: Proceedings of the Workshop on Temporal and Spatial Reasoning, IJCAI '93.

Asher, N.–Sablayrolles, P. 1994. A typology and discourse semantics for motion verbs and spatial PPs in French. To appear in: Journal of Semantics.

Aurnague, M.–Vieu, L. 1993. Towards a formal representation of space in language: a commonsense reasoning approach. In: Proceedings of IJCAI '93.

Asher, N.–Wada, H. 1989. A computational account of syntactic, semantic and discourse principles for anaphora resolution. In: Journal of Semantics.

Boons, J.P. 1985. Préliminaires à la classification des verbes locatifs: les compléments de lieu, leurs critères, leurs valeurs aspectuelles. In: Linguisticae Investigationes 9: 195–267.

Boons, J.P. 1987. La notion semantique de deplacement dans une classification syntaxique des verbes locatifs. In: Langue Française 76: 5–40.

Clarke, B.L. 1981. A calculus of individuals based on connection. In: Notre Dame Journal of formal logic 22: 204–18.

Dervillez-Bastuji, J. 1982. Structure des relations spatiales dans quelques langues naturelles—introduction à une théorie sémantique. Library Droz, Geneva–Paris.

Guillet, A. 1990. Une classification des verbes transitifs locatifs. Ph.D. thesis. Université Paris 7, LADL.

Hayes, E.M. 1989. On defining motion verbs and spatial prepositions. Bericht B61, SFB 314: VITRA, Fachbereich 10—Informatik IV. Saarbrücken, Germany.

Kamp, H. 1979. Events, instants and temporal reference. In: Bäuerle, R.-Egli, U.-von Stechow, A. (eds): Semantics from different points of view, 376-417. Springer Verlag, Berlin.

Lascarides, A.-Asher, N. 1993. Temporal interpretation, discourse relations, and commonsense entailment. In: Linguistics and Philosophy 16: 437-93.

Lamiroy, B. 1987. Les verbes de mouvement, emplois figures et extensions métaphoriques. In: Langue française 76: 41-58.

Laur, D. 1991. Sémantique du déplacement et de la localisation en français: une étude des verbes, des prépositions et de leurs relations dans la phrase simple. Ph.D. thesis. Université Toulouse le Mirail, Toulouse, France.

Laur, D. 1993. La relation entre le verbe et la préposition dans la sémantique du déplacement. In: Language, La couleur des prépositions 47-67.

Lesniewski, S. 1927-31. O podstawach matematyki. In: Przegląd filozoficzny No. 30-40. Partial translation by G. Kalinowski in: Lesniewski, S. 1989. Sur les fondements de la mathématique. Hermes, Paris.

Maienborn, C. 1992. Kompakte Strukturen: Direktionale PPn und nicht-lokale Verben. In Bosch, P.-Gerstl, P. (eds): Sprachtheoretische Grundlagen für die Computerlinguistik, 131-45. DFG Sonderforschungsbereich 340, Stuttgart, Bericht Nr. 30.

Wunderlich, D. 1991. How do prepositional phrases fit into compositional syntax and semantics? In: Linguistics 29: 591-621.

Addresses of the authors: Nicholas Asher
IRIT, University Paul Sabatier
118 route de Narbonne
F-31062 Toulouse Cedex
France
e-mail: asher@irit.fr

Pierre Sablayrolles IRIT, University Paul Sabatier
118 route de Narbonne
F-31062 Toulouse Cedex
France
e-mail: sablay@irit.fr

Acta Linguistica Hungarica, Vol. 42 (3–4), pp. 253–270 (1994)

AN INFINITE NUMBER OF MONKEYS*

DAVID BEAVER

1. Introduction

Life is short. There is not enough time to explain everything. As speakers, or writers, we are forced to make assumptions. It is common to be advised to fix in one's mind a picture of the audience, that is, to make an advance decision as to what the audience can be expected to know. Often, especially given limitations of time for speaking or space for writing, one is forced to take much for granted. As a result, cases of presupposition failure, the situation occurring when the speaker or writer takes for granted something of which the hearer or reader is not previously aware, are surely common. Somehow, hearers and readers cope, and usually without complaining.

The author, in most *genres*, assumes that the text will be read linearly, and further assumes, optimistically, that readers will gather information throughout the reading process. So what the author has is not a fixed picture of the common ground with the intended readership, but a rather rough cut and idealised movie of how this common ground should develop. Each frame in the movie approximates what is common between relevant aspects of (1) the author's beliefs at time of writing, and (2) the readers' beliefs as they reach some point in the text.

At risk of straining the cinematic metaphor somewhat, it could be said that the text itself is analogous to a script, but with detailed screenplay and directorial instructions omitted. Barring a major scientific breakthrough, the corresponding film will never be put on general release, so precisely how the writer intends the information state of idealised readers to evolve as they read is never made public in all its technicolour glory.

Shortly (in section **3**), I will discuss how readers' information states do evolve, as they construct their own movies on the basis of the script. In directing their own films readers second-guess the intentions of the original writer-

* This is a shorter version of a manuscript currently in preparation under the working title "A Plea for Common Sense".

1216–8076/94/$ 5.00

director. But to understand how readers work out the writer-director's intentions, it is firstly necessary to know more of the craft of the writer-director. In the coming section I will elaborate on how, working from an assumption as to initial conditions, the author envisages the evolution of the common ground. It will be helpful to adopt some of the formalism of recent dynamic semantics. I will build particularly on ideas of Stalnaker (1974), Karttunen (1974) and Heim (1983), and use formal techniques related to those discussed by Groenendijk–Stokhof (1991), and Veltman (1991).

2. The writer's view of the common ground

I will now present what will here be called Presupposition Logic, a simple propositional system with a dynamic semantics and dynamic notion of semantic entailment. Further discussion and motivation can be found in Beaver (1992, 1993b, 1994, 1995).

Presupposition Logic provides a model of how a speaker or author envisages the common ground evolving. This evolution is iterative, since the common ground at any instant provides the context in which a given chunk of text is interpreted, and it is the effect of this interpretation which determines what the common ground will be going into interpretation of the next chunk. It is no longer controversial to assert that the interpretation process relies on such iteration, but there remains some question as to the course-grainedness of the iteration. For instance in the work of Gazdar (1979) (and also related proposals such as Mercer's 1992) it is whole sentences which produce a change in the context of interpretation. However in Karttunen–Heim style treatments of presupposition such as that introduced in this paper, as well as in treatments of anaphora due to Heim (1982), Kamp (1981) and Groenendijk–Stokhof (1991), a finer grained iteration is involved, with sub-sentential constituents producing their own effects on local contexts of interpretation.

We begin by assuming some set of atomic proposition symbols. A model is a pair $\langle W, I\rangle$, where W is a set of worlds and I is an interpretation function mapping each atomic proposition symbol to a subset of W. The Context Change Potential (to borrow Heim's terminology) of a formula ϕ written $[\![\phi]\!]$ is a set of pairs of input and output contexts, where a context is the writer's view of the common ground. Following Stalnaker, a context is thought of as a set of possible worlds, the set containing all and only those worlds compatible with the information supposed to be common. I will write $\sigma[\![\phi]\!]\tau$ to indicate that the pair $\langle\sigma, \tau\rangle$ is in the $[\![\phi]\!]$ relation, i.e. an input context σ updated with the information in ϕ can produce an output context τ.

Defining contexts as subsets of W introduces a natural lattice structure with union and intersection as meet and join, and this lattice provides to an intuitive model-theoretic characterisation of the amount of information in a given context. A minimal context (with respect to some model) may be defined as the set of all worlds (in that model): this is the state of blissful ignorance in which no information about the world is available. Similarly a maximally informative non-contradictory context would be a singleton set: the available information rules out all except one world. However, it is possible to add even more information to such a context, and in case this information contradicts previous information we will arrive at a context containing no worlds, a truly maximal but contradictory context.

Definition D1, below, gives an update semantics for the language of propositional logic. The first clause says that the result of updating a context with an atomic proposition is an output containing only those worlds in the input which are in the extension of the proposition. A context can be updated with a conjunction of two formulae (the second clause) just in case it can firstly be updated with the left conjunct to produce an intermediary context (υ in the definition), and this context can be updated with the right-hand conjunct to produce the final output (τ). The third clause says that a context can be updated with the negation of a formula just in case there is some state that can be obtained by updating the context with the negated formula itself, in which case the result of updating with the whole formula is the set of worlds in the input which are not present in the update with the negated formula. In other words, the effect of updating with the negation of a formula is to remove all information compatible with the formula. The final two clauses define implication and disjunction by (carefully selected) standard equivalences.

Definition D1 (Update Semantics for Propositional Logic) *For all models $\mathcal{M} = \langle W, I\rangle$ and information states σ, τ, the relation $[\![.]\!]^{\mathcal{M}}$ (superscript omitted where unambiguous) is given recursively by:*

(1) $\sigma[\![p_{\text{atomic}}]\!]\tau$ *iff* $\tau = \{w \in \sigma \mid w \in I(p)\}$

(2) $\sigma[\![\phi \wedge \psi]\!]\tau$ *iff* $\exists \upsilon\ \sigma[\![\phi]\!]\upsilon[\![\psi]\!]\tau$

(3) $\sigma[\![\neg\phi]\!]\tau$ *iff* $\exists \upsilon\ \sigma[\![\phi]\!]\upsilon \wedge \tau = \sigma \backslash \upsilon$

(4) $\sigma[\![\phi \rightarrow \psi]\!]\tau$ *iff* $\sigma[\![\neg(\phi \wedge (\neg\psi))]\!]\tau$

(5) $\sigma[\![\phi \vee \psi]\!]\tau$ *iff* $\sigma[\![\neg(\neg\phi \wedge \neg\psi)]\!]\tau$

A context σ satisfies a formula ϕ ($\sigma \models \phi$) if updating adds no new information, producing an output identical to the input. One formula entails another ($\phi \models \psi$) if any update with the first produces a context in which the second is satisfied:

Definition D2 (Satisfaction and Dynamic Entailment)

$$\begin{array}{lll} \sigma \models \phi & \textit{iff} & \sigma[\![\phi]\!]\sigma \\ \phi \models \psi & \textit{iff} & \forall \sigma, \tau \;\; (\sigma[\![\phi]\!]\tau \Rightarrow \tau \models \psi) \end{array}$$

Over the standard propositional language this notion of entailment is classical. However, we will now extend the language with a presupposition operator ∂: the resulting logic will be non-classical. For example, commuting conjunctions will no longer uniformly preserve validity. The intuition behind the following definition is that a formula $\partial\phi$ ("the presupposition that ϕ") places a constraint on the input context, only allowing update to continue if the presupposed proposition is already satisfied.

Definition D3 (Presupposition Logic) *Presupposition Logic is defined over the language of Propositional Logic with an extra unary operator ∂. It has the dynamic notion of semantic entailment above, and semantics consisting of the update semantics for atomic propositions and standard connectives combined with the following interpretation for ∂-formulae:*

$$\sigma[\![\partial\phi]\!]\tau \quad \textit{iff} \quad \sigma \models \phi$$

As an example the sentence "Mary realises that John is sleepy" might be said to correspond to a formula in Presupposition Logic of the form $\partial p \wedge q$, where p is atomic proposition that John is sleepy, and q is an atomic proposition that Mary has come to believe that John is sleepy. In this paper I will not be concerned with the details of how such logical forms may be derived from natural language, or with the question of whether it is reasonable to use a representation in which presuppositions are divided explicitly from assertions using the ∂-operator. However, both these issues are dealt with elsewhere: see Beaver (1993a, 1993b, 1995) where two sorted versions of classical type theory are used to provide a Presupposition Logic style semantics for a fragment of natural language.

Suppose a formula contains a presuppositional sub-formula. What will the presuppositions of the whole formula be? This is the presupposition projection problem of Langendoen–Savin (1971), except applied to Presupposition Logic rather than natural language. A formula $\partial\phi$, "the presupposition that ϕ", defines an update if and only if ϕ is satisfied, so it is natural to say that in general a formula presupposes all those formulae that must be satisfied by the input context in order for there to be an update. We say that a context σ admits a formula ϕ (written $\sigma \triangleright \phi$) if and only if it is possible to update σ with ϕ, this being a formalisation of Karttunen's notion of admittance in Karttunen (1974). In that case one formula ϕ presupposes another ψ (written $\phi \gg \psi$) just in case every context that admits the first satisfies the second. Note that admittance provides a counterpart to so-called **presupposition failure**, what happens when updating cannot continue because presuppositions are not satisfied.

Definition D4 (Admits ($\triangleright$) and Presupposes ($\gg$))

$$
\begin{array}{rll}
\sigma \triangleright \psi & \textit{iff} & \exists\tau\sigma[\![\phi]\!]\tau \\
\phi \gg \psi & \textit{iff} & \forall\sigma\ \sigma \triangleright \phi \Rightarrow \sigma \models \psi
\end{array}
$$

It is clear that the definitions for the semantics of the ∂-operator and the meta-logical $\gg$ relation are closely related. For example we have that for any ϕ, $\partial\phi \gg \phi$. Indeed, $\gg$ could have been equivalently defined in terms of ∂, defining $\phi \gg \psi$ *iff* for some χ, $[\![\phi]\!] = [\![\partial\psi \wedge \chi]\!]$. We can now study projection in Presupposition Logic. As detailed in F1, the system behaves just as anyone familiar with Karttunen's 1974 system would expect. In particular, note that a formula may fail to carry a presupposition of one of its component subformulae, but instead carry a logically weaker conditionalised variant.

Fact F1

If $\phi \gg \psi$ then:

$$
\begin{array}{rcl}
\neg\phi & \gg & \psi \\
\phi \wedge \chi & \gg & \psi \\
\phi \rightarrow \chi & \gg & \psi \\
\phi \vee \chi & \gg & \psi \\
\chi \wedge \phi & \gg & \chi \rightarrow \psi \\
\chi \rightarrow \phi & \gg & \chi \rightarrow \psi \\
\chi \vee \phi & \gg & (\neg\chi) \rightarrow \psi
\end{array}
$$

Before going on, I should perhaps point out a shortcoming of Presupposition Logic *qua* model of natural language. Whilst Presupposition Logic is like natural language in allowing both assertion and presupposition of information, it is very unlike natural language in that presuppositions in Presupposition Logic seem so pointless. Granted, they have some interesting logical properties, as F1 shows, but further they seem somewhat gratuitous. In natural languages, the phenomenon of presupposition is not an arbitrarily occurring feature, but an essential part of the way language works. Presuppositions are not simply poor cousins that hang around whilst assertions do the real work of communication. Rather, presuppositions can serve an essential function in determining what actually is asserted. For instance, a definite description "the king of France", would ordinarily (say in a history lesson) be used to identify an object so that something can be predicated of it, and the presuppositions associated with the description, of existence and maybe uniqueness of the object, are not incidental, but essential to the functioning of the description. In Presupposition Logic, presuppositions do not play such important roles. However, we will see in the next section that presuppositions like those of Presupposition Logic could still play a part in the communication process. Although such presuppositions do not directly help to establish what is asserted, as is the case in natural language, they can at least provide a murky reflection of what has been taken for granted by the writer, that is, they can give readers an insight into what the writer took to be the initial common ground.

3. The naive reader

Readers who updated their own information state according to principles like those behind the semantics of Presupposition Logic would be stymied whenever some information was presupposed which they did not have. In this situation the reader would lack any further means of updating. So Presupposition Logic, as it stands, does not provide a good model of the evolution of the information state of a hearer or reader. But suppose that you only had a Presupposition Logic-like semantics to help you understand a text. How would you use that semantics to glean information?

If an infinite number of monkeys with typewriters were given time, some of them might produce this text. What if an infinite group of monkeys schooled in the writer-director approach to writing described above, were given some rhetorical goal (for instance, to produce my contribution to the Noszvaj symposium) but no description of the intended audience? Not knowing the audience, they might choose the initial common ground randomly, although after that each monkey's view of the common ground at a particular point in the text would be fully determined by what they typed. Now although as a reader you do not know what initial common ground has been assumed when you read a text, you can reason for any particular choice of initial conditions how the common ground would evolve. And this leads me to a suggestion for how a dynamic semantics like that given for Presupposition Logic could be used to understand a text without presupposition failure being problematic.

Begin the reading process by imagining an infinite number of monkeys (or as many as you can manage) with an infinite number of assumed initial common grounds, your goal being to find out which monkey wrote the text. As you read, separately update each of these contexts. At various stages presuppositional constructions will be encountered, and these are what sort out the wheat from the chaff. For whenever something is presupposed which has not been explicitly introduced earlier in the text, a number of monkeys drop out of contention, and there remain only those monkeys for which the assumed common ground corresponding to that point in the text satisfies the presupposition. In general, this process may not tell you exactly which monkey was responsible, but it will at least limit the options, and it will simultaneously tell you quite a lot about what information you were intended to have after reading the text.

Let us consider a simple example concerning a language with only two proposition letters p and q. Assume a model with four possible worlds, w^{pq}, w^{p}, w^{q} and w: in the first both p and q hold, in the second only p, in the third only q, and in the fourth, neither. Monkeys can be identified with the set of worlds which they associate with the common ground, so each monkey's assumed context corresponds to a subset of $\{\mathrm{w}^{pq}, \mathrm{w}^{p}, \mathrm{w}^{q}, \mathrm{w}\}$. Thus there are 16 relevantly different monkeys, each of which we can represent using four digit binary numbers. Let the units, twoes, fours and eights digits correspond to worlds w, w^{q}, w^{p} and w^{pq} respectively, such that a given digit is 1 if and only if the corresponding world is in the assumed common ground. For instance, monkey 1010 assumes that the common ground contains worlds w^{pq} and w^{q}.

You read a text $\partial p \wedge q$, and wish to know which monkey wrote it. It is easily verified, from the semantics presented in the previous section, that the only pairs in $[\![\partial p \wedge q]\!]$ are $\langle 0000, 0000\rangle$, $\langle 0100, 0000\rangle$, $\langle 1000, 1000\rangle$ and $\langle 1100, 1000\rangle$. In words: an inconsistent context (0000) remains inconsistent after update with the formula; a context containing only world w^{p} can be updated, but this world is a non-q world, and is eliminated by the update with q, producing an inconsistent context; an update with an input context containing only world w^{pq} returns the input as output; a context containing worlds w^{pq} and w^{p}, thus a context where p is established but q is still uncertain, can be updated with the effect of removing world w^p, the non-q world, to produce a context containing only world w^{pq}. No other contexts can be updated. For instance it is not possible to update the context 1011, which satisfies neither p or q, because the presence of non-p worlds causes presupposition failure. So the semantics of presupposition logic tells us that one of the monkeys 0000, 0100, 1000 or 1100 must have written the text.[1]

The many-monkeys strategy can easily be formalised in terms of the semantics of Presupposition Logic. A reader's information state is identified with a set of contexts—I will use the term **information set**—and is thus a subset of the powerset of worlds.[2] A state can be updated with a formula by updating

[1] One could use simple Gricean principles to reason further: the maxim of informativeness tells us that some information should have been transmitted. Thus neither 0000 (a state from which no further update is possible) nor 1000 would have been written by pragmatically sophisticated monkeys, since they both produce identity updates. If the maxim of *quality* is taken to rule out updates with contradictory information, then 0000 is ruled out once again (because it contradicts all information), and 0100 is also ruled out, because update then leads to the inconsistent state. Thus the only remaining monkey would be 1100, who assumed p and used the formula to communicate that q. In spite of the appeal of this sort of approach to formalisation of Grice's work (an approach which is perhaps reminiscent of that adopted by van der Sandt (1992)) I will pursue it no further here.

[2] Observe that two assumptions have been implicitly made, firstly that the author and the reader have agreed upon a common set of atomic proposition letters, and secondly that they use identical models (up to isomorphism). Without this latter assumption, there would be no question of the reader speculating as to which worlds were in the author's assumed common ground, for there would be no sense in which the reader could ever know anything of the author's worlds or *vice versa*. One might ask whether these assumptions, in effect that author and reader use the same vocabulary and lexical semantics, are significant. Adding uncertainty about the vocabulary would complicate technically, but it seems that not much would be gained. For there is no technical reason why the chosen set of atomic proposition symbols could not be much larger than the vocabulary that could conceivably be used, say the set of all finite sequences of Roman characters and Arabic numbers. The assumption of a common model is more metaphysically puzzling, but also, I will argue, need not be problematic.

There is a natural ordering amongst the models of Presupposition Logic, some being more constrained than others. An unconstrained model would be such that for every world w and proposition symbol p, there is another world w' which is in the extension of exactly the same proposition symbols as w, except for being in the extension of p only if w is not, and *vice versa*. A term model constructed by identifying each subset of the proposition symbols with a world, and identifying the extension of each proposition symbol with the set of worlds containing that symbol, would in this sense be completely unconstrained. A model would be more constrained if there were certain sets of proposition symbols such that no world was in the extension of all and only those symbols. Meaning postulates produce constrained models. For current purposes we may suppose that the model is unconstrained. This means only that we have not stated any special conventions about how proposition symbols should be interpreted, and we are, in a sense, in a state of mutual ignorance. In choosing an unconstrained model, we assume only that there are no special assumptions. To me, at least, this does not seem worrying.

each of the member contexts separately, so producing the following definition of the update of a state I with a formula ϕ:

Definition D5 (Updating Information Sets)

$$I + \phi = \{\tau \mid \exists\sigma \in I\ \sigma[\![\phi]\!]\tau\}$$

By definition, let us say that an information set satisfies a formula only if its member contexts satisfy the formula:

Definition D6 (Satisfaction by an Information Set)

$$I \models \phi \quad \textit{iff} \quad \forall\sigma \in I\ \sigma \models \phi$$

The earlier notion of entailment could easily be defined in terms of the new notion of information, as the following fact demonstrates. (Here $\mathcal{P}(\mathcal{W})$ is the powerset of the set of worlds W.)

Fact F2

$$\begin{aligned} \phi \models \psi \quad & \textit{iff} \quad (\mathcal{P}(\mathcal{W}) + \phi) \models \psi \\ & \textit{iff} \quad \forall I\ (I + \phi) \models \psi \end{aligned}$$

To return to the earlier example, we have the following update: $\{0000, \ldots, 1111\} + \partial p \wedge q = \{0000, 1000\}$. The reader who accepts the veridicality of the text can then learn both p and q, since both these propositions are satisfied in $\{0000, 1000\}$. However, this is not to say that there is no longer a difference between presupposed and asserted formulae. Consider update with the negation of the previous example formula. We have that $\{0000, \ldots, 1111\} + \neg(\partial p \wedge q) = \{0000, 0100\}$, and whilst the result state $\{0000, 0100\}$ still satisfies p, it no longer satisfies q: in fact it satisfies $\neg q$. Thus the presupposition (that p) is effectively projected out of the negation, whilst the assertion, as would be expected, is not.

4. The sophisticated reader

The naive reader might imagine an infinite number of monkeys, and use only information from the text to help find out which monkey is the author. But other information is available, if not of an absolute character. We cannot initially say of any given proposal as to the assumed common ground that it is impossible, and to this extent it is necessary to consider all possibilities. But we can say that some proposals are relatively more plausible than others. The sophisticated reader considers what assumption the author is **likely** to have made as to the initial common ground.

I cannot say how we decide what the author is likely to have assumed. One strategy would be to guess that the world about which the author was trying to communicate obeys similar, or identical, general principles as your own world does, and that the author had a similar knowledge of those principles. That is, you could guess that the author's common sense knowledge is similar to your own. In fact the model I will now propose will be of sufficient generality that the particular method of guessing at the author's assumptions will not be important. The assumptions of the author, whatever they are, determine a Presupposition Logic context, a set of worlds. A reader's knowledge of which assumptions are most plausible determines an ordering over these contexts, what I will call a *plausibility ordering*. A plausibility ordering relative to some model is a reflexive, transitive binary relation over a subset of the powerset of the set of worlds.[3] For an ordering π, $\sigma \geq_\pi \tau$ is written for $\langle\sigma,\tau\rangle \in \pi$, and $\sigma >_\pi \tau$ is taken to mean that both $\sigma \geq_\pi \tau$ and $\tau \not\geq_\pi \sigma$. An ordering π can be updated with a new formula by considering every pair in the ordering, and updating each element of the pair separately according to the principles of Presupposition Logic. The following definition is obtained:

Definition D7 (Updating Plausibility Orderings)

$$\pi + \phi \quad = \quad \{\langle\sigma,\tau\rangle \mid \exists\langle\sigma',\tau'\rangle \in \pi \; \sigma'[\![\phi]\!]\sigma \;\wedge\; \tau'[\![\phi]\!]\tau\}$$

Under this definition, certain contexts may drop out of contention in the update process, just as with the naive updating process considered earlier. An example may clarify. Suppose that $[\![\phi]\!] = \{\langle\sigma,\sigma'\rangle,\langle\upsilon,\upsilon'\rangle\}$, and that we wished to update the ordering $\pi = \{\langle\sigma,\sigma\rangle,\langle\tau,\tau\rangle,\langle\upsilon,\upsilon\rangle,\langle\sigma,\tau\rangle,\langle\sigma,\upsilon\rangle,\langle\tau,\upsilon\rangle\}$ with ϕ. Returning to the earlier metaphor, the reader is considering three different **movies** that the writer-director might have intended, and at the current point in the text the candidates for the correct frame are σ, τ and υ,

[3] László Kálmán, in a talk presented at the University of Amsterdam, also proposed a treatment of presupposition involving default knowledge encoded using orderings. One difference is that he envisaged a model in which possible worlds were ordered rather than a model in which sets of worlds are ordered. However, I think we are in agreement on the basic issues, in particular on the fact that some encoding of default knowledge must be used in the treatment of presupposition.

with a plausibility ordering $\sigma \geq_\pi \tau \geq_\pi \upsilon$. The reader should now verify that $\pi + \phi = \{\langle\sigma',\sigma'\rangle, \langle\upsilon',\upsilon'\rangle, \langle\sigma,\upsilon\rangle\}$. Observe that since τ cannot be updated with ϕ, there is no next frame from the film containing τ in the new ordering, and we are left with only two candidate films, with current frames σ' and υ' ordered $\sigma' \geq_{\pi+\phi} \upsilon'$. So the fact that frames σ and υ were in a certain ordering relation means that the next frames in those films are in the corresponding ordering relation after update. A more sophisticated model would perhaps allow juggling of orderings in the update process, to allow for what Grice-style conversational analysis might tell us about the author's knowledge and intentions.

Before considering how we might make use of plausibility orderings, let us see how the earlier notion of entailment could be defined in terms of them. The *domain* of an ordering π, written $\star\pi$, can be defined as the set of contexts which are at least as plausible as themselves in the ordering, and this allows retrieval from a plausibility ordering of a corresponding information set. This in turn permits the definition of a notion of satisfaction of a formula by a plausibility ordering in terms of the earlier notion of satisfaction by an information set:

Definition D8 (Domain of an ordering and 'Ordinary' Satisfaction w.r.t. an ordering)

$$\begin{array}{rcl} \star\pi & = & \{\sigma \mid \sigma \geq_\pi \sigma\} \\ \pi \models \phi & \textit{iff} & \star\pi \models \phi \end{array}$$

Given such a notion of satisfaction, it should be clear that it would be straightforward to define a notion of entailment equivalent to that given earlier. However, it is also possible to define alternative notions of entailment relative to any given plausibility ordering. Let us say that the set of **preferred contexts** in an ordering π, written $\uparrow \pi$, is the set of all contexts which are at least as plausible as any context in the ordering. Then we can say that an ordering π **preferentially satisfies** a formula ϕ, written $\pi \triangleright \phi$, if the set of preferred contexts in π satisfies ϕ. Preferential satisfaction is a weaker notion than satisfaction, in that an ordering may preferentially satisfy more formulae than it satisfies. We may now say that a formula ϕ preferentially entails a formula ψ relative to an ordering π, written $\phi \triangleright_\pi \psi$, if updating π with ϕ produces an ordering which preferentially satisfies ψ. Here are the formal definitions:

Definition D9 (Preferential Satisfaction and Entailment)

$$\begin{array}{rcl} \uparrow \pi & = & \{\sigma \mid \forall \tau \in \star\pi\ \sigma \geq \tau\} \\ \pi \rhd \phi & \mathit{iff} & \uparrow \pi \models \phi \\ \phi \rhd_{\pi} \psi & \mathit{iff} & \pi + \phi \rhd \psi \end{array}$$

In the case of a trivial ordering consisting of the cross-product of the powerset of worlds $\pi_0 = \mathcal{P} \times \mathcal{P}$, for which every set of worlds is at least as plausible as every other set of worlds, this notion collapses into the earlier entailment:

Fact F3

$$\phi \models \psi \quad \mathit{iff} \quad \phi \rhd_{\pi_0} \psi$$

Consider the following pair of sentences:

(1) *If Jane takes a bath, Bill will be annoyed that there is no more hot water.*

(2) *If Jane wants a bath, Bill will be annoyed that there is no more hot water.*

An utterance of (1) does not suggest to me that there actually is no more hot water, but only that if Jane takes a bath, there will be no more hot water. On the other hand, (2) suggests strongly that there is no more hot water. Put another way, (1) is compatible with the standard CCP prediction of a conditional reading, but (2) is not.

The current theory will predict the contrast provided the following plausibility assumptions hold:

- At least one alternative in which it is established that there is no hot water is more plausible than all alternatives in which it is not known whether there is hot water, but in which it is known that if Jane wants a bath then there will be no hot water.
- An alternative in which it is not known whether or not there is hot water but in which it is established that if Jane has a bath then there will be no more hot water must be at least as plausible as all alternatives where it is definitely established that there is no hot water.

The general question is, why would it be reasonable to expect plausibility orderings to have such properties? My answer to this question is on the one hand both simple and obvious, and on the other hand both awkward to implement and incompatible with many contemporary theories of presupposition. Many linguists will surely find it unpalatable. The answer is: common sense.

Let me expand on this. The contrast between (1) and (2) results from our ability to find a common-sensical explanation of the lack of hot water in terms of somebody having taken a bath, but in our inability to fully explain a lack of hot water in terms of somebody simply wanting a bath. The simple assumption that there is a finite amount of relevant hot water—about a bathful—is sufficient to allow justification of there being no more hot water in situations where Jane has just taken a bath. However, the same simple assumption would not suffice in the case of (2), and a number of other assumptions would be needed, such as the assumption that if Jane wants a bath then she will definitely take one. Thus it is the relative plausibility of assumptions not explicitly mentioned in the text of the example sentences that determines what is implicated.

Let us see how some of this analysis of (1) and (2) may be crudely formalised. For expositional purposes, I will ignore many obviously relevant issues, such as temporal connections between antecedent and consequent clauses in the conditionals. Let us represent "Jane takes a bath" as JTB, "Jane wants a bath" as JWB, "there is no hot water" as NHW, and "Bill will be annoyed that there is no more hot water" as ∂NHW $\wedge$ BA. Now suppose that our common sense knowledge of the relative plausibility of different assumptions is encoded in a plausibility ordering π. The two conditions required of π may be formalised as follows:

- $\exists\sigma \in \pi\ \sigma \models \text{NHW}$ and
  $\forall\tau \in \pi$ if $(\tau \not\models \text{NHW}$ and $\tau \models \text{JWB} \rightarrow \text{NHW})$ then $\sigma >_\pi \tau$
- $\exists\sigma \in \pi\ \sigma \not\models \text{NHW}$ and $\sigma \models \text{JTB} \rightarrow \text{NHW}$ and
  $\forall\tau \in \pi$ if $\tau \models \text{NHW}$ then $\sigma \geq_\pi \tau$

If π conforms to these requirements, then we have the following preferential entailments:

$$\text{JWB} \rightarrow (\partial\text{NHW} \wedge \text{BA}) \quad \rhd_\pi \quad \text{NHW}$$
$$\text{JTB} \rightarrow (\partial\text{NHW} \wedge \text{BA}) \quad \not\rhd_\pi \quad \text{NHW}$$
$$\text{JTB} \rightarrow (\partial\text{NHW} \wedge \text{BA}) \quad \rhd_\pi \quad \text{JTB} \rightarrow \text{NHW}$$

In other words, with respect to π, (1) preferentially entails that there is no more hot water, and (2) preferentially entails not there is no more hot water, but that if Jane takes a bath then there is no more hot water.

5. Discussion

The theory developed in this paper can be thought of as an attempt at a formal characterisation of what Lewis (1979) called **accommodation**, most authors since having followed his terminology, with the exception of Seuren (1985), who refers to an apparently similar process which he terms **backward suppletion**. But in its current form the model differs markedly from existing proposals, in that most writers have taken accommodation to be some sort of repair strategy, something that happens when the interpretation process goes wrong. Lewis seems to picture accommodation as a covert adjustment of what he calls the **conversational score**, a sort of creative accounting needed to make conversational ends meet. Van der Sandt's accommodation, to take a more recent example, is a sophisticated cut-and-paste operation on discourse representation structures. Thus accommodation has been viewed as an essentially non-monotonic operation, overwriting our previous record of what had happened in a discourse to fit with new demands. The view espoused in this paper has been quite the contrary. Accommodation is seen as a **monotonic** operation, in the sense that it does not replace or destructively revise our information about a speaker or author, but further instantiates our knowledge, reducing the range of possibilities for what the speaker was assuming.

Yet this monotonicity is not crucial, it is merely symptomatic of the differences between the approach I have espoused and others. The main claim I wish to make is that when we accommodate, we look not only at the record of what has been said, but also behind what has been said, and consider explicitly what the author might have intended and what the author might have expected. Once we explicitly include in our model the reader's uncertainty about the authors assumptions, and if we place no restrictions on the reasoning power of the reader, as a side effect our model will tend to predict that the reader's information state develops monotonically.

So we see two extremes. Firstly there is the Lewisian model where the reader keeps a score of what has been said, but occasionally retrospectively alters that score to fit the facts. Secondly there is the model presented here, where the reader's energy is devoted to second-guessing the author's assumptions, and the reader's own score is thus a higher order entity recording all the possibilities for what the author takes the score to be. The reality must lie between these extremes: reasoning about the author is expensive from a processing point of view, so readers are forced to take short cuts. This means that readers cannot simultaneously consider all the alternative contexts the author might be assuming, and will occasionally be forced to back-track and search for alternative contexts that fit the facts better than those under consideration. Yet it remains the case that some sort of reasoning about the writer's assumptions is involved, and that a model of accommodation is unlikely to be successful unless it takes into account such reasoning.

I am very much in agreement with Richmond Thomason, who's notion of accommodation is yet more sophisticated than the one I have proposed here, since it takes into account not only the beliefs, but also the communicative goals of other agents: "Concentrating on accommodation means shifting to reconstructed reasoning that underlies utterances. And it suggests that certain reasoning processes, such as intention recognition and cooperation are central. Successful accommodation requires that we first recognize someone's intention to achieve a goal, and then establish goals of our own that will assist in achieving this goal" (Thomason 1990). For the sake of concreteness, I have attempted to escape the need for this sophistication by restricting my attention to a very simple presuppositional language, and by concentrating on the understanding of texts, where the scope for development of cooperative plans is naturally limited.

The model I am proposing is certainly not the first theory of presupposition to take into account the beliefs of other agents. Indeed, in a certain sense Gazdar's theory (Gazdar 1979) and some of its descendants (e.g. van der Sandt 1988, Mercer 1992, Blok 1993) go further in this direction than I have. What I particularly like about these theories is their utilisation of epistemic logics allowing for explicit reasoning about agents beliefs. However, in all of these theories the type of reasoning about what has been assumed by a speaker or writer is somewhat limited. Given that a presupposition trigger has been used, the only question is whether the triggered presupposition is or is not believed by the speaker. On the other hand in the current model a much richer connection between the triggered presuppositions and the assumptions of the speaker or writer is allowed.

I will now briefly discuss one point of confusion in the account I have offered. For (1), there is the possibility either of accommodating that *if Jane takes a bath there is no hot water* or accommodating that *there is no hot water* simpliciter. The weaker conditional appears to be preferred. On the other hand, the treatment of (2) seems to involve the reverse situation, with a logically stronger (i.e. non-conditionalised) proposition being accommodated.

How can it be that sometimes a logically stronger proposition is considered more plausible than a logically weaker one, and sometimes *vice versa*?[4] This apparent enigma arises from a misidentification of the accommodation process with the addition of single propositions, and a misidentification of the ordering over epistemic alternatives as an ordering over propositions. The plausibility ordering should not, however, be thought of as an ordering over propositions, but over closed sets of propositions, i.e. **theories**. Thus I am not committed to any claim about the conditional *if Jane takes a bath there is no hot water* being more plausible than the proposition *there is no hot water*, and I am not committed to *there is no hot water* being more plausible than *if Jane wants a bath there is no hot water*. Rather, I would claim that there must be a relevant closed theory containing the first conditional which is more plausibly taken to be the common ground than every relevant closed theory containing the simple proposition. Similarly, there must be a theory containing the simple proposition which is more plausible than every theory which contains the second conditional but not the simple proposition. There is no reason to assume that any relation of inclusion (the **theory** level counterpart of propositional entailment) holds between these various theories.

So contexts could profitably be thought of as logically closed theories, and what is accommodated does not hang on the relative plausibility of alternative propositions, but on the relative plausibility of alternative theories. This particular point would become much clearer if the mechanisms described in this paper were presented more proof theoretically. In Karttunen's earliest presentation of a context change model, that in Karttunen (1974), contexts are defined as sets of propositions. It would be natural to recast the current model in this way. A similarly constructive alternative would be to use DRS's as contexts, in which case one might end up with a picture close to that of van der Sandt. And rather than supposing that a reader simultaneously keeps in mind all the different possible contexts, we could model a reader's information states in terms of a single context σ together with some representation of common-sense knowledge about the world. Then the update operation would become essentially abductive. Given that the speaker has uttered ϕ, the reader's task would be to establish the minimal most plausible set of assumptions the speaker could be making which contains σ and allows update with ϕ to produce the new context. Such developments will have to await future research. But it can be seen that the formal theory presented in in this paper

[4] This objection, that plausibility appears oddly disconnected from entailment, was first made to me by Henk Zeevat. The same objection can be found in Geurts (Ms).

is a more model-theoretic counterpart of the sort of abductive picture of the interpretation process for which Jerry Hobbs (see e.g. Hobbs–Stickel–Appelt–Martin 1993) has long been arguing.[5]

References

Beaver, D. 1992. The kinematics of presupposition. Proceedings of the Eighth Amsterdam Colloquium. ILLC, University of Amsterdam.

Beaver, D. 1993a. Kinematic Montague Grammar. In: Kamp, H. (ed.): DYANA-2 deliverable R2.2a. University of Amsterdam.

Beaver, D. 1993b. What comes first in dynamic semantics. ILLC report LP-93-15. University of Amsterdam.

Beaver, D. 1994. When variables don't vary enough. In: Harvey, M.–Santelmann, L. (eds): SALT 4. Cornell.

Beaver, D. 1995. Presupposition and assertion in dynamic semantics. Ph.D. dissertation. University of Edinburgh.

Beaver, D. Presupposition. To appear in: van Benthem, J.–ter Meulen, A. (eds): The Handbook of Logic and Linguistics. Elsevier, Amsterdam.

Blok, P. 1993. The Interpretation of focus: an epistemic approach to pragmatics. Doctoral dissertation. University of Groningen.

Gazdar, G. 1979. Pragmatics: implicature, presupposition and logical form. Academic Press, New York.

Groenendijk, J.–Stokhof, M. 1991. Dynamic predicate logic. In: Linguistics and Philosophy 14: 39–100.

Geurts, B., Presupposing. Ms. University of Osnabrück.

Heim, I. 1982. On the semantics of definite and indefinite noun phrases. Ph.D. dissertation. University of Amherst.

Heim, I. 1983. On the projection problem for presuppositions. In: WCCFL '83.

Heim, I. 1990. Presupposition projection. In: Reader for the Nijmegen Workshop on Presupposition, Lexical Meaning, and Discourse Processes. University of Nijmegen.

Hobbs, J.–Stickel, M.–Appelt, D.–Martin, P. 1993. Interpretation as abduction. In: Artificial Intelligence 63: 69–142.

Kamp, H. 1981. A theory of truth and semantic representation. In: Groenendijk, J.–Janssen, T.–Stokhof, M. (eds): Formal Methods in the Study of Language.

Karttunen, L. 1974. Presuppositions and Linguistic Context. In: Theoretical Linguistics 1.

[5] One could also, I think, draw a profitable analogy between the treatment of presupposition in this paper and the common-sense driven analyses of discourse and temporal relations in the work of Lascarides–Asher (1993). One might take various temporal expressions (e.g. tensed verbs) as presupposing some temporal reference with accommodation then being necessary to build the most plausible link between previously introduced temporal discourse referents and the new reference point.

Langendoen, D.–Savin, H. 1971. The projection problem for presuppositions. In Fillmore, Ch.J.–Langendoen, D. (eds): Studies in Linguistic Semantics. Holt, Reinhardt & Winston, New York.

Lascarides, A.–Asher, N. 1993. Temporal interpretation, discourse relations and commonsense entailment. In: Lingustics and Philosophy 16: 437–95.

Lewis, D. 1979. Scorekeeping in a language game. In: Bäuerle, R. *et al.* (eds): Semantics from Different Points of View. Springer Verlag, Berlin.

Mercer, R. 1992. Default logic: towards a common logical semantics for presupposition and entailment. In: Journal of Semantics 9/3.

van der Sandt, R. 1988. Context and Presupposition. Croom Helm, London.

van der Sandt, R. 1992. Presupposition projection as anaphora resolution. In: Journal of Semantics 9/4.

Seuren, P. 1985. Discourse Semantics. Basil Blackwell, Oxford.

Soames, S. 1982. How presuppositions are inherited: a solution to the projection problem. In: Linguistic Inquiry 13: 483–545.

Stalnaker, R. 1974. Pragmatic presuppositions. In: Munitz, D.K.–Unger, D.K. (eds): Semantics and Philosophy.

Thomason, R. 1990. Accommodation, meaning and implicature: interdisciplinary foundations for pragmatics'. In: Cohen, P.–Morgan, J.–Pollack, E. (eds): Intensions in Communication, 325–63. Bradford Books, MIT Press, Cambridge, Mass.

Veltman, F. 1991. Defaults in update semantics. In: DYANA deliverable R2.5.C (also: to appear in: Journal of Philosophical Logic).

Zeevat, H. 1992. Presupposition in update semantics. In: Journal of Semantics 9/4.

Address of the author: David Beaver
Centre for Cognitive Science
University of Edinburgh
and
Department of Philosophy
University of Amsterdam
e-mail: dib@illc.uva.nl

Acta Linguistica Hungarica, Vol. 42 (3-4), pp. 271-282 (1994)

FLEXIBLY DISAMBIGUATING METONYMY

LEON VERSCHUUR

0. Introduction

In this paper I will show how to disambiguate metonymic expressions while doing justice to the fact that often several of their senses can be used simultaneously. I will first give an overview of some empirical properties of metonymy, and sketch the classical pragmatic account of Nunberg (1978, 1979) of these properties. That theory did not postulate ambiguity, which leads, as I will show, to problems. Then I will present a semantic, type-theoretic account which reconciles the assumption of ambiguity with the possibility of simultaneous occurrence of senses.

1. The ambiguity of metonymy

A rather uncontroversial of metonymy would be: the phenomenon that a lexical expression, often a noun, is applicable to different categories of objects, related by certain language-independent concepts and their compositions. Examples, stemming mainly from Pustejovsky (1994b) and Nunberg (1978), are:

(1) Form–Content (EXPRESS)
  (a) Mary believes the book.
  (b) Mary sold the book to John.

(2) Place–Inhabitant (LOCATED IN)
  (a) John traveled to New York.
  (b) New York kicked the mayor out of office.

(3) Product–Producer (PRODUCE)
  (a) The newspaper decided to fire its editor.
  (b) John spilled coffee on the newspaper.

(4) Container–Containee (CONTAIN)
  (a) Mary broke the bottle.
  (b) The baby finished the bottle.

(5) Process–Result (RESULT IN)
  (a) The company's merger with Honda will begin next fall.
  (b) The merger will produce cars.

Now these are largely intuitive characterisations of the sentences in terms of categories and their metonymic relations, insofar that it is difficult to arrive at a principled classification of these notions. Note for instance that the categories are not necessarily disjoint, since a person or an institute may be both an inhabitant and a producer, and a physical object may be both a form and a product. Further, one might consider for instance the relation Form–Content to be a case of Container–Containee. I have not tried to develop an exhaustive and principled classification of metonymic relations either, but will simply assume that it can be made, in any case for limited domains of discourse. Assumed these metonymyic notions then, we can observe the following generalisations about metonymic expressions.

First, metonymic nouns are according to the standard criteria for ambiguity not *vague*, but *ambiguous*. The familiar truth-conditional test yields that the truth of a sentence can depend on the metonymic sense chosen:

(6) Clara Schumann loved Brahms. (Product–Producer)

(7) He went to this school. (Place–Inhabitant)

An utterance of (6) could in principle be true if the Producer sense is taken, but false if the Product sense is understood. To contrast this with vagueness, one cannot judge *a student has arrived* false in case a female student has arrived on the ground that we chose a presumed 'male student' sense—even though we can maintain that *student* is vague between these two interpretations, since other expressions may specify its meaning to one of them (see for instance Verschuur 1993). Similarly, (7) can in principle at the same time be true if the Building-sense is chosen, and false in the case of the Institute-sense.

Second, also the VP-anaphora test yields that metonymic senses are distinct meanings:

(8) ?John bought a newspaper, and so did Robert Maxwell. (Product–Producer)

This has no crossed (Product–Producer) reading, at best if interpreted as a pun.

Now whereas metonymy is an instance of ambiguity, it is important to contrast it with plain **homonymy**. Besides the existence of productive relations between senses, metonymy can be discriminated from homonymy by the possibility of selecting several senses simultaneously:

(9) A thick dissertation has been refuted. (Form–Content)

(10) The bottle which he threw away had tasted disgusting. (Container–Containee)

(11) The newspaper has decided to change its format. (Producer–Product)

(12) A school has burnt down. It is bankrupt now. (Place–Inhabitant)

Of these, in (9) and (10) the metonymic noun is first modified by an expression selecting one sense, after which the verb phrase selects another one. In (11) and (12) first the main verb selects one sense, and then an NP-anaphor for which the immediate context selects a different sense, still manages to refer back to the individuals introduced by using the original senses. These discourse cases exemplify what we may call an NP-anaphor test, the cases (9) and (10) a modifier-noun test.

2. A pragmatic and a semantic explanation

Nunberg has in the cited works proposed to explain the ambiguity as well as the possible cooccurrence of senses as follows. There would be just one sense listed in the lexicon (which one is in a certain sense undetermined); after semantic interpretation, a pragmatic apparatus of 'referring functions' (modelling the metonymic relations) and their compositions take us from the individuals in this one extension to individuals of other categories, thereby yielding a pragmatic acceptability of the utterance. This explains, given that the referring functions would be characterised exhaustively,

- the productivity of metonymic concepts
- the truth-conditional differences between senses
- the cases of sense cooccurrence.

But this pragmatic approach cannot explain

- the impossibility of VP-anaphora
- how to get a decidable procedure recognizing the acceptable propositions.

As for the first item, VP-anaphora using an interpretation different from the one present in the antecedent (such as in (8)) would certainly have been possible if the assumption of only one lexical meaning would have been correct. Finally, it has not been elaborated which decidable procedure would derive the pragmatic acceptability on the basis of the referring functions and their compositions.

The present approach assumes that all senses correspond to different, typed predicates ("generated" by Nunbergian metonymic functions); and that parsing is constrained by checking of semantic types in a system of type assignment to λ-terms, including the rules for intersective conjunction (see Lambek 1961; Kanazawa 1992; Morrill 1990). This will explain

- the productivity of metonymic concepts
- truth-conditional differences between senses
- the cases of sense cooccurrence
- the impossibility of VP-anaphora
- how to get a decidable procedure recognizing the acceptable propositions.

The first two items here are self-evident; the impossibility of VP-anaphora with using different sense follows from the standard assumption that the meaning of a VP-anaphor and of its antecedent are the same, or have at least a common generalisation. So I will concentrate further on the possibility of simultaneous selections of senses, as well on framing a decidable procedure to be used in parsing. Now as for the latter, there are two possibilities:

a) constrain an arbitrary grammar by semantic type deduction in an undirected typed λ-Calculus of a Gentzen sequent format, for mechanical proof construction (see Barendregt–Hemerik (1990) for a type-theoretic exposition and Gallier (1986) for sequent formalisations in general);

b) use a Categorial Calculus with directed semantic types, where the directionality constrains the syntactic possibilities, and the atomic sorts within them are motivated semantically (see Benthem 1991).

In the following sections, I will elaborate the second option, which is the more perspicuous and less complex of the two.

Before turning to that, I will briefly mention however another (related) kind of account, namely the one using Type Coercion rules licensing various kinds of type change. These have been elaborated within the framework of Generative Lexicon Theory (GLT), being developed by Pustejovsky, and have been applied there to logical polysemy in general (among which metonymy), see Pustejovsky (1994a, 1994b), Buitelaar–Mineur (1994); also they have been elaborated, more specifically and exclusively for the cases of metonymy, in work of Dölling (1994a, 1994b). In general this approach of Type Coercion differs from the one in this paper that our type shifting rules are not so much motivated by the empirical phenomena, but are simply the standard introduction and elimination rules (in our Gentzen formalisation: the right and left rules) of the proof theory expressing the usual meaning of the type constructors; only the very choice of the type constructors (in particular for the conjunction) is motivated empirically. Nevertheless, a closer comparison between the results here and in these two frameworks is still needed.

3. A categorial calculus accounting for metonymy

The special features a categorial grammar accounting for metonymy needs to have include the following:

- many-sortedness instead of the traditional *entity* type, sorts modelling conceptual categories such as *form*, *content*, *inhabitant*, *place*, ...
- transitive closure of sorts under conjunction, with certain conditions on both construction and interpretation of complex sorts
- rules for introducing and eliminating intersective conjunction types
- rules for introducing and eliminating (restricted) universal quantification types.

Since instantiation of a sort variable will be restricted however to the finite number of sorts, the quantification rules are eliminable; their only purpose is to escape from having to assign to for instance determiner representations a large number of types—instead, one polymorphic type will suffice.

Conjoined sorts may not be construed at random, as one might expect; for any conjoined sort there has to be a metonymic function $EXPRESS$, $LOCATED\ IN, \ldots$ which maps each inhabitant of one conjunct to an inhabitant of the other (or to an inhabitant of an *element* of the other, in case this conjunct is complex itself).

Further there is the restriction on interpretation that each n-tuple in the domain of a conjunction contains an element of which all others depend as the values of metonymic functions. For instance, $[\![institute \wedge place \wedge product]\!]$ includes

$$< the\ Ford\ factory,\ the\ Ford\ location,\ the\ Ford\ cars >$$

but not

$$< the\ Ford\ factory,\ the\ Opel\ location,\ the\ Toyota\ cars >$$

In other words, we can refer to complex individuals, but they must be coherent and may not be scattered.

Before I will give the calculus and a sample lexicon, a few words about the interpretation of what I will call "directed semantic types", of which an example would be $e\backslash t$ instead of $NP\backslash S$. The interpretation with respect to the domain of expressions is completely analogous to the standard one of syntactic types as you can see below; the only deviation is that I have chosen semantic symbols for the atomic types instead of syntactic ones, the reason being that as a consequence of our categorisation into sorts, there is no longer a one-to-one mapping from syntactic to semantic types. In particular the semantic types for nouns will now be far more diverse than the single syntactic type CN (common noun). (An obvious disadvantage of this choice is that *purely* syntactic differences between expressions become difficult to represent now, so in the end the picture would possibly have to be refined.)

Interpretation with respect to a domain of expressions (if E is the set of lexical expressions closed under concatenation)

$$[\![A/B]\!]_{Syn}] = \{x \in E | \forall y \in [\![B]\!]_{Syn}, xy \in [\![A]\!]_{Syn}\}$$
$$[\![A\backslash B]\!]_{Syn}] = \{x \in E | \forall y \in [\![A]\!]_{Syn}, yx \in [\![B]\!]_{Syn}\}$$
$$[\![A \wedge B]\!]_{Syn}] = \{x \in E | x \in [\![A]\!]_{Syn}, x \in [\![B]\!]_{Syn}\}$$
$$[\![A \bullet B]\!]_{Syn}] = \{xy \in E | x \in [\![A]\!]_{Syn}, y \in [\![B]\!]_{Syn}\}$$

Note that whereas the $\bullet$ denotes the familiar product type constructor of the original Lambek calculus (Lambek 1958), the $\wedge$ stands for the intersective conjunction (mentioned by Lambek in Lambek (1961) and applied more recently by Morrill in Morrill (1990)). It denotes a set of single expressions belonging to two types at once (and not a set of *pairs* of expressions like the product does).

Interpretation with respect to a domain of entities

$$[\![A/B]\!]_{Sem}] = [\![A]\!]_{Sem}^{[\![B]\!]_{Sem}}$$
$$[\![A\backslash B]\!]_{Sem}] = [\![B]\!]_{Sem}^{[\![A]\!]_{Sem}}$$
$$[\![A \wedge B]\!]_{Sem}] = [\![A \bullet B]\!]_{Sem} = [\![A]\!]_{Sem} \times [\![B]\!]_{Sem}$$

Here the interpretation of product and intersective conjunction is the same, only this means that for intersection a single expression denotes an ordered pair, and for the product the two expressions together denotes an ordered pair. By the way, in my application the single expressions for ordered pairs are only *hypothesised* during deductions, as we will see. For the directed implications finally, their semantic interpretation is identical to that of the standard (undirected) arrow type.

I will notate the calculus in the style of Morrill, where a symbol between brackets behind a sequence Γ denotes an occurrence of a certain sequence within Γ, and with $\delta'\beta$ for functional application of δ to β. The implication rules are essentially the ones of Lambek (1958). Intersective conjunction was expressed in the form of two axioms in Lambek (1961), but here also in Morrill style Left and Right rules.

Lambek calculus with intersective conjunction and restricted quantification

$$\alpha : X \vdash \alpha : X[Ax]$$

$$\frac{\Delta \vdash \beta : Y \qquad \Gamma(\delta'\beta : X) \vdash S}{\Gamma(\delta : X/Y, \Delta) \vdash S}[/L] \qquad \frac{\Gamma, x : X \vdash \beta : Y}{\Gamma \vdash \lambda x.\beta : Y/X}[/R]$$

$$\frac{\Delta \vdash \beta : Y \qquad \Gamma(\delta'\beta : X) \vdash S}{\Gamma(\Delta, \delta : Y\backslash X) \vdash S}[\backslash L] \qquad \frac{x : X, \Gamma \vdash \beta : Y}{\Gamma \vdash \lambda x.\beta : X\backslash Y}[\backslash R]$$

$$\frac{\Gamma(\pi_1\alpha : X) \vdash S}{\Gamma(\alpha : X \wedge Y) \vdash S}[\wedge L] \quad \frac{\Gamma(\pi_2\alpha : Y) \vdash S}{\Gamma(\alpha : X \wedge Y) \vdash S}[\wedge L] \quad \frac{\Gamma \vdash \alpha : X \quad \Gamma \vdash \beta : Y}{\Gamma \vdash <\alpha, \beta>: X \wedge Y}[\wedge R]$$

$$\frac{\Gamma(\alpha : X[\mathcal{X} := Sort]) \vdash S}{\Gamma(\alpha : \forall\mathcal{X}.X) \vdash S}[\forall L] \qquad \frac{\Gamma \vdash \alpha : X}{\Gamma \vdash \alpha : \forall\mathcal{X}.X}, \mathcal{X} \notin FV(\Gamma)[\forall R]$$

According to the left rules for conjunction, reasoning upwards from conclusion to premise, one may choose only one of the two projections of the pair, corresponding to one of the two types of the expression in question. The quantification rules are in the variant without universal abstraction (by Λ), because it is not necessary here to express explicitly the dependence of (sub)terms on types.

4. Application by a lexical grammar

Given this calculus, the correct analysis of for instance

(9) A thick dissertation has been refuted.

is obtained by assuming the following lexicon

A sample lexicon

a$\rightsquigarrow$ $\lambda PQ.\exists x(Px \wedge Qx) : \forall X.(t/(X\backslash t))/(X\backslash t)$
thick$\rightsquigarrow$ $thick : (f\backslash t)/(f\backslash t)$
dissertation$\rightsquigarrow$ $dissertation_1 : f\backslash t$
dissertation$\rightsquigarrow$ $dissertation_2 : c\backslash t$
has been refuted$\rightsquigarrow$ $refuted : c\backslash t$

in which f abbreviates the *form* sort and c abbreviates the *content* sort. I have given the second translation of 'dissertation' $dissertation_2$ only in order to emphasize the ambiguity of representation which is assumed; this representation will not be used at all in the deduction of the sentence, since it cannot be selected by *thick*.

$$a : (t/((f \wedge c)\backslash t))/((f \wedge c)\backslash t)\ thick : (f\backslash t)/(f\backslash t)\ dis_1 : f\backslash t\ ref : c\backslash t \vdash a(\lambda A\ thick(dis_1)(\pi_1(A)))(\lambda B\ ref(\pi_2(B))) : t \quad [\mathbf{L}/]$$

$$dis_1 : f\backslash t \vdash dis_1 : f\backslash t \quad [\mathbf{Ax}]$$

$$a : (t/((f \wedge c)\backslash t))/((f \wedge c)\backslash t)\ thick(dis_1) : f\backslash t\ ref : c\backslash t \vdash a(\lambda A\ thick(dis_1)(\pi_1(A)))(\lambda B\ ref(\pi_2(B))) : t \quad [\mathbf{L}/]$$

$$thick(dis_1) : f\backslash t \vdash \lambda A\ thick(dis_1)(\pi_1(A)) : (f \wedge c)\backslash t \quad [\mathbf{R}\backslash]$$

$$A : f \wedge c\ thick(dis_1) : f\backslash t \vdash thick(dis_1)(\pi_1(A)) : t \quad [\mathbf{L}\wedge]$$

$$\pi_1(A) : f\ thick(dis_1) : f\backslash t \vdash thick(dis_1)(\pi_1(A)) : t \quad [\mathbf{L}\backslash]$$

$$\pi_1(A) : f \vdash \pi_1(A) : f \quad [\mathbf{Ax}]$$

$$thick(dis_1)(\pi_1(A)) : t \vdash thick(dis_1)(\pi_1(A)) : t \quad [\mathbf{Ax}]$$

$$a(\lambda A\ thick(dis_1)(\pi_1(A))) : t/((f \wedge c)\backslash t)\ ref : c\backslash t \vdash a(\lambda A\ thick(dis_1)(\pi_1(A)))(\lambda B\ ref(\pi_2(B))) : t \quad [\mathbf{L}/]$$

$$ref : c\backslash t \vdash \lambda B\ ref(\pi_2(B)) : (f \wedge c)\backslash t \quad [\mathbf{R}\backslash]$$

$$B : f \wedge c\ ref : c\backslash t \vdash ref(\pi_2(B)) : t \quad [\mathbf{L}\wedge]$$

$$\pi_2(B) : c\ ref : c\backslash t \vdash ref(\pi_2(B)) : t \quad [\mathbf{L}\backslash]$$

$$\pi_2(B) : c \vdash \pi_2(B) : c \quad [\mathbf{Ax}]$$

$$ref(\pi_2(B)) : t \vdash ref(\pi_2(B)) : t \quad [\mathbf{Ax}]$$

$$a(\lambda A\ thick(dis_1)(\pi_1(A)))(\lambda B\ ref(\pi_2(B))) : t \vdash a(\lambda A\ thick(dis_1)(\pi_1(A)))(\lambda B\ ref(\pi_2(B))) : t \quad [\mathbf{Ax}]$$

In this derivation the rule applications have been printed upside down, with cumulative indentation of each premise or pair of premises. I have omitted the $\forall L$ application with which the deduction begins, instantiating the sort variable X in the determiner term to $f \wedge c$. Further I have abbreviated the determiner term by a.

The essence of the explanation of cooccurrent senses is that whenever two predicates are true of two individuals, here $thick(dissertation_1)$ and $refuted$, one can deduce that related predicates are true of the pair of these individuals (in addition such pairs are assumed to *exist* according to the restricted sort instantiation of $\forall L$). For instance, according to the fourth line, if $thick(dis_1)$ is applicable to a form, then of course $\lambda x.thick(dis_1)(\pi_1 x)$ is applicable to a form–content pair. In the sixth line from below one can see an analogous sequent involving $refuted$. The term in the succedent of the conclusion—originally offered as a variable to which a t was assigned—is eventually instantiated by a term consisting of the determiner applied to both these complex, derived predicates. This term abbreviates

$$\lambda PQ.\exists x(Px \wedge Qx)(\lambda A\ thick(dissertation_1)(\pi_1(A)))(\lambda B\ refuted(\pi_2(B)))$$

which β-reduces to

$$\exists x_{f\wedge c}(thick(dissertation_1)(\pi_1 x) \wedge refuted(\pi_2 x))$$

That $\pi_2 x$ denotes an individual metonymically related to $\pi_1 x$ (and so in the extension of $dissertation_2$) follows from the restrictive interpretation of conjunction types: the ordered pair $< [\![\pi_1 x]\!], [\![\pi_2 x]\!] >$, being in the domain of $form \wedge content$, must contain an element of which the other depends as the value of a metonymic function applied to it—in this case the function $EXPRESSION$. So $[\![\pi_2 x]\!]$ is a content expressed by $[\![\pi_1 x]\!]$, and not an arbitrary content.

Also analyses of texts with an NP-anaphor using a predicate meaning different from the metonym meaning within its antecedent can in principle be derived, like (11)–(12). This can be worked out using the dynamic λ-representations of Muskens (1990), because the semantic theory of agreement of anaphors in Verschuur (1993), which embeds a type theory with intersection, union, and complement constructors into this dynamic semantics, still makes the right prediction if we add conjunction types. The theory says that individual terms may be coindexed if there is a type assignment from which all types of them are deducible. It would lead for (12) to the assumption of a type *location* $\wedge$ *institute* assigned to the pair of coindexed terms associated with 'a school' respectively 'it'. We will show this on another occasion.

References

Barendregt, H. – Hemerik, K. 1990. Types in Lambda Calculi and Programming Languages. In: Jones, N. (ed.): European Symposium on Programming, 1–36. Springer, Berlin.

van Benthem, J. 1991. Language in Action. North-Holland, Amsterdam.

Buitelaar, P. – Mineur, A. 1994. Coercion and compositionality in Categorial Grammar. In: Dekker, P. – Stokhof, M. (eds): Proceedings of the 9th Amsterdam Colloquium, 175–88. ILLC, Amsterdam.

Dölling, J. 1994a. Semantic sorts and systematic ambiguity. In: Robering, K. (ed.): Sorten, Typen, Typenfreiheit. Klassifikationsprobleme linguistischer Einheiten. Working Papers in Linguistics, Bd. 30. Technical University, Berlin.

Dölling, J. 1994b. Bedeutungsvariationen und lexikalische Nominalstruktur. In: KONVENS 94: Das Lexikon in der Sprachverarbeitung. Springer, Berlin – Vienna – New York.

Gallier, J. 1986. Logic for Computer Science. Foundations of Automatic Theorem Proving. Harper and Row Publishers, New York.

Girard, J.-Y. 1989. Proofs and Types. Cambridge University Press, Cambridge.

Kanazawa, M. 1992. The Lambek Calculus enriched with additional connectives. In: Journal of Logic, Language, and Information 1/2: 141–71.

Lambek, J. 1958. The mathematics of sentence structure. In: American Mathematical Monthly 65: 154–69.

Lambek, J. 1961. On the calculus of syntactic types. In: Jacobson, R. (ed.): Structure of Language and its Mathematical Aspects. American Mathematical Society.

Morrill, G. 1990. Grammar and logical types. In: Stokhof, M. – Torenvliet, L. (eds): Proceedings of the Seventh Amsterdam Colloquium, 429–50. ITLI, Amsterdam.

Muskens, R. 1990. Anaphora and the logic of change. In: van Eijck, J. (ed.): Logics in AI, Proceedings of JELIA '90. Lecture Notes in AI Nr. 478. Springer-Verlag, Berlin.

Nunberg, G. 1978. The Pragmatics of Reference. Indiana University Linguistics Club, Bloomington.

Nunberg, G. 1979. The non-uniqueness of semantic solutions: Polysemy. In: Linguistics and Philosophy 3. Linguistics Club, Bloomington.

Pustejovsky, J. 1994a. Semantic typing and degrees of polymorphism. In: Vide, M. (ed.): Current Issues in Mathematical Linguistics: 221–38. Elsevier, Amsterdam.

Pustejovsky, J. 1994b. Linguistic constraints on type coercion. In: Saint-Dizier, P.–Vegas, E. (eds): Computational Lexical Semantics. Cambridge University Press, Cambridge.

Verschuur, L. 1993. Semantic agreement of anaphors and extended type theory. In: Kathol, A.–Bernstein, M. (eds): Proceedings of the Eastern States Conference On Linguistics '93. Cornell University Press.

Address of the author: Leon Verschuur
Faculteit der Letteren (TKI)
Tilburg University
Warandelaan 2
Postbus 90153
5000 LE Tilburg
The Netherlands
e-mail: L.F.Verschuur@kub.nl

Acta Linguistica Hungarica, Vol. 42 (3-4), pp. 283-303 (1994)

AN INDEXICAL THEORY OF SCOPE

DONKA F. FARKAS

1. Introduction

In this paper I propose a theory of scope that is less structure driven than the traditional approach. The traditional view of scope is structural in the sense that the relative scope of two expressions is taken to be determined by their relative position at some level where hierarchical relations are encoded. More precisely, in this view, e_1 is in the scope of e_2 iff e_2 **commands** e_1 at the appropriate structural level. I take the term **command** in its generic sense here, meaning 'higher than' and leave the details of how to define its domain unspecified for now. Common to all varieties of command is that it is defined at the sentence level.

The structural approach to scope has immediate repercussions for syntax under the assumption that semantic structure is read off of syntactic structure; the original motivation for the syntactic level of LF was precisely the assumption that command relations in syntax determine semantic command relations and therefore scope. Under this assumption, e_1 is in the scope of e_2 iff e_2 commands e_1 at LF.

My aim here is to propose a theory in which the relative scope of two expressions is a matter of possible dependencies between indices, seen as Kaplan-style parameters of evaluation. The proposal builds on insights concerning temporal reference in Enç (1986), and is close in spirit to work in Discourse Representation Theory (DRT). The main difference between the indexical theory of scope to be presented here and the DRT approach found in Kamp–Reyle (1993) is that scope relations in DRT are encoded structurally at the level of Discourse Representation Structure: different scopal relation induce DRSs that differ structurally. In the view presented here the difference is one of content rather than structure.

Structural considerations will turn out to be relevant here too, but the theory differs from the traditional one in the following two respects: (i) scope is seen as essentially discoursal; there is no attempt at reducing discoursal

1216-8076/94/$ 5.00

effects to sentential ones as in the accommodation approach of Roberts (1989) and Poesio–Zucchi (1992); (ii) structural considerations at both the syntactic and the semantic level are seen as underdetermining scopal relations: structure determines when an expression **may** be in the scope of another, but not when it **must** be in its scope. Because of space limitations, discoursal issues will remain almost entirely outside the scope of this paper.

After outlining basic assumptions in section **2**, I turn to presenting the proposal in section **3**. Section 4 shows how the indexical theory presented here accounts for the scopal properties of indefinite and distributive noun phrases, and mentions some of the problems that remain open.

2. Scope of what with respect to what

I am concerned here with the scope of noun phrases (or DPs) with respect to intensional operators (modals, the conditional operator), intensional predicates (such as *believe*, *dream*), and other 'quantificational' DPs, i.e., DPs whose D contributes a proportional (strong) quantifier. I restrict my attention to DPs whose contribution to LF, the representation of semantic structure, includes a subscripted variable x_n, and a descriptive content, DC_n, in the form of a predicative expression on x_n. In addition, quantificational DPs, i.e., DPs whose D is a quantifier such as *every*, *each* or *most*, induce a tripartite quantificational structure in which the quantifier is given by the determiner, and where the variable and the DC occur in the Restrictor. In this case the quantificational force of the DP is determined by the quantifier it contributes. I follow DRT and File Change Semantics in assuming that the quantificational force of non-quantificational DPs depends on their position in the semantic structure. I will come back below to the question of exactly how expressions containing such free variables are to be interpreted.

(i) Scope of the DC

A descriptive condition DC_n constrains the assignment of values to the variable x_n to assignments where the value in question meets the descriptive condition. One aspect of the issue of noun phrase scope concerns the question of determining the world in which this condition has to be met. Thus, (1a) is ambiguous depending on whether the variable introduced by the underlined noun phrase is to be a friend of mine in w, the world of evaluation of the whole sentence, or in w_J, the world according to John, introduced by the matrix predicate *believe*; (1b) can be interpreted either with the DC evaluated with respect to w, or with the DC evaluated with respect to the worlds introduced and quantified over by the modal.

(1) (a) *John believes that* a friend of mine *is a crook.*
(b) A friend of mine *might be a crook.*

When the world of evaluation is w, the DC is said to have wide scope with respect to the predicate and the modal; when the world of evaluation is that introduced by the predicate or the modal the DC will be said to have narrow scope with respect to them.

A QR-based account of DC scope has the noun phrase occur in a position commanding the intensional predicate or modal at LF in the case of the wide scope reading, and it has the predicate or modal command the noun phrase at LF in the narrow scope reading. The problem of the scope of the DC of a noun phrase with respect to intensional predicates or operators is reminiscent of the issue of the temporal interpretation of the DC discussed in Enç (1986). Enç shows that the temporal reference of the DC of a noun phrase is in principle independent of the temporal reference of the main predication of which the noun phrase is an argument. The account of the scope of the DC to be proposed here builds on this similarity.

(ii) Scope of the variable

I turn now to another aspect of scope that arises in connection with quantification, and that I will call here the scope of the variable. The issue concerns the question of whether there is co-variance of values assigned to some variable v with values assigned to some other variable v'. The expression that contributes v is said to be within the scope of the expression that contributes v' in the dependency case and outside its scope otherwise.

I am assuming a standard tripartite quantificational structure at the level of LF, where the Restrictor introduces a set of 'cases', assumed to be non-empty, which form the domain of quantification. A quantificational expression whose quantifier is Q is true iff Q-many ways of making the Restrictor true are also ways of making the Nuclear Scope (NS) true. What the 'cases' in the domain of quantification are depends on the type of quantificational expression involved. The domain of quantification is a set of worlds in modal sentences and afactual (or non-indicative) conditionals such as those exemplified in (2). The domain of quantification is a set of situations within a single world in adverbial quantification and factual (or indicative) conditionals of the type exemplified in (3), and it is a set of individuals in nominal quantification, exemplified in (4).

(2) (a) *John may be sick.*
(b) *If John were sick, he'd be in bed.*

(3) (a) *Sometimes, when John is hungry he is grouchy.*
(b) *If John is hungry he is grouchy.*

(4) *Every student left.*

In order to verify a quantificational claim one has to identify the set of relevant cases based on the information provided by the Restrictor, and then one has to check whether the possibly complex property expressed by the NS is true of the appropriate number of cases. In (2b), (3b) and (4) there is an overt restrictive phrase: the antecedent in (2b) and (3b), the *when* clause in (3a), and the semantic content of the subject noun phrase in (4). The context may further restrict the relevant set of cases to a salient subset of the cases identified by the overt restriction. In (2a) there is no overt restrictive phrase. What the relevant set of cases is here is determined by the interpretation of the modal and contextual factors. I will concentrate below on cases where there is an overt restrictive phrase.

Following DRT, I assume that both instances of existential closure proposed in Heim are dispensed with. Text level existential closure is rendered superfluous, as suggested in Heim (1982, Ch. 3), by the requirement that the text be true in the model with respect to which it is interpreted. If truth conditions are stated in terms of assignment functions, this reduces to the requirement that there be an assignment function that satisfies the conditions imposed by the DRS or logical form (LF) constructed on the basis of the text. NS-level existential closure is rendered superfluous by the requirement that the NS be true of the appropriate number of cases introduced by the Restrictor.

I am assuming that expressions are interpreted with respect to a model M such that $M = <W, U, V, F>$, such that W is a set of worlds, U is a set of individuals, V is a set of valuation functions assigning intensions to constants, and F is the set of assignment functions that assign intensions to variables.

It will be important in what follows to be able to keep track of individuals across worlds. This can be done either by allowing the domains of worlds to overlap or by enriching the model with counterpart relations that connect individuals across worlds. I adopt the former option though nothing crucial depends on this choice. Below I briefly discuss the relevant features of nominal and modal quantification. For the sake of brevity issues concerning quantification over situations will be left outside the scope of this paper.

2.1. Nominal quantification

The 'cases' that form the domain of quantification are the individuals in w (or the individuals in some contextually salient situation in w) that meet the DC condition of the quantificational DP. A way of making such a case true amounts to choosing an evaluation function that assigns to the variable contributed by the quantificational DP a value that meets the DC condition of the DP. The logical form of sentences involving nominal quantification is as in (5),

(5) Restrictor Qx_n Nuclear Scope

where x_n is the variable contributed by the quantificational DP.[1] Truth conditions for an expression of this form are given in (6).

(6) A quantificational LF of the form in (5) is true in w wrt M iff there is an assignment function f such that there are Q-many assignment functions f_R that extend it and that verify the LF in the Restrictor, such that each f_R has an extension f_{NS} which verifies the NS.

The notion of a function being the extension of another, defined as in (7), is needed to keep track of assignments of mentioned variables.

(7) Extension of f
An assignment function f_e is an extension of an assignment function f iff f_e agrees with f on all mentioned variables in f.

A function is said to verify an LF iff it meets the conditions of that LF; a variable x is mentioned in f iff f verifies some LF in which x appears.

The LF of (4) is (8):

(8) Restrictor:	$\forall x_1$	Nuclear Scope:
x_1		*leave* $'(x_1)$
student $'(x_1)$		

[1] Whether the quantifier unselectively binds any or all other variables in the Restrictor is a question that I leave open here. Heim (1982) assumes unselective binding. In traditional analyses that treat indefinites as quantificational, the assumption is that such variables are bound by a narrow scope existential quantifier. To implement the latter solution in a Heimian framework one would need to introduce a Restrictor-level existential closure operation. The choice between these alternatives has repercussions with respect to the 'proportion problem'.

The quantifier binds x_1 because it is contributed by the DP that introduces x_1. The truth conditions of the LF in (8) are given in (9).

(9) The LF in (8) is true in w with respect to M iff there is some assignment function f with the following property: every assignment function f_R which extends f such that $f_R(x_1) \in V(student')$ at w has the property of having an extension f_{NS} such that $f_{NS}(x_1) \in V(leave')$ at w.

(Since the NS here has no new variable relative to the Restrictor, $f_R = f_{NS}$.) The evaluation function f will be referred to as the *base* evaluation function. The interpretation of expressions involving nominal quantification involves, besides the base function, the set of functions f_R, which are extensions of f that verify the Restrictor, and the set of functions f_{NS}, which extend the functions in f_R and verify the Nuclear Scope.

The evaluation of x_1 above has to be done relative to f_R because x_1 is the variable contributed by a quantificational DP. In examples that exhibit scope ambiguities non-quantificational DPs contribute variables in the Restrictor or the Nuclear Scope. I will restrict my attention here to the latter case, exemplified in (10):

(10) *Every student speaks an Indo-European language.*

In (10), the material in the Restrictor is the variable contributed by the subject noun phrase and its DC. The cases introduced by the Restrictor, which form the domain of quantification, are the individuals in the context which satisfy the Restrictor. The NS contains the material contributed by the indefinite, as well as the main predication. The sentence is true just in case the NS is true with respect to each case that satisfies the Restrictor.

The two interpretations of (10) that are known as the 'wide scope' and the 'narrow scope' reading of the indefinite result from the option of evaluating the variable by the base assignment function f (wide scope) or by the set of functions f_{NS} (narrow scope). In structural accounts of scope this difference corresponds to a difference in the structural position of the variable contributed by the indefinite. In Quantifier Raising (QR) accounts, the indefinite noun phrase is raised so as to command the quantificational noun phrase at LF to give the wide scope reading, while in the case of the narrow scope reading QR results in a configuration where the quantificational noun phrase commands the indefinite. The analysis in NP-preposing terms in Heim (1982) achieves the same result. In DRT the variable contributed by the indefinite is added to the main DRS box to give the wide scope reading, and to the NS box to give the narrow scope reading.

2.2. Modal quantification

Following Kratzer (1977) and Kratzer (1981), modal sentences such as (2) are assumed to involve quantification over worlds. The 'cases' that form the domain of quantification are a subset of the worlds W. In simple modal sentences, such as (2a) the worlds in the relevant subset are those in which the contextually supplied propositions in the modal base are true. In modal conditionals, the worlds in the domain of quantification are the set of worlds W_R such that the worlds in this set are close to the base world w, and are such that the proposition expressed by the antecedent is true in them. I will restrict my attention to conditionals here. The LF of a modal conditional is a quantificational structure with the LF contributed by the antecedent, LF_a, in the Restrictor, and the LF contributed by the consequent, LF_b, in the NS. The quantifier binds the worlds in W_R:

(11) Restrictor: $\forall \mathrm{w}_R$ Nuclear Scope:
LF_a LF_q

A way of making the Restrictor true here amounts to the choice of a world $w_R \in W_R$. The truth conditions will be as in (12).

(12) An LF of the form in (11) is true in w wrt M iff there is an assignment function f such that for every w_R that is close to w and which has the property that f verifies LF_a at w_R, f verifies LF_q at w_R.

(I will not be concerned here with the 'closeness' requirement.) The issue that arises now is the choice of world parameters for particular subparts of LF_a and LF_q. The available choices are the base world, w, and the worlds w_R in W_R. Evaluating a subpart of the conditional with respect to the base world gives rise to the wide scope reading of that expression with respect to the conditional; evaluation of such a subpart with respect to w_R gives rise to the narrow scope reading of the expression with respect to the conditional. In QR-based accounts these choices correlate with different structural positions at LF: wide scope interpretation correlates with a position commanding the conditional, and narrow scope correlates with a position commanded by the conditional. In DRT the wide scope interpretation correlates with material entered in the main DRS box, and the narrow scope interpretation correlates with material entered in the antecedent or the consequent boxes. To exemplify, consider (13):

(13) *If someone were here she'd help.*

Under the wide scope reading of the indefinite in the antecedent, one has to find a value for the variable it contributes within the domain of the base world w, such that in all worlds w_R in which that person is here, she helps. Under the narrow scope reading of the indefinite, each world w_R such that there is an individual in that world who is here is such that that individual helps.

The ambiguity illustrated here concerned the choice of variable. In examples such as (14),

(14) *If a friend of mine were here she'd help.*

the question of what world the DC must be met in arises as well. The question there is whether the value assinged to the variable contributed by the indefinite has to be a friend of mine in w or w_R.

Note that all cases of narrow sope variables discussed here involved interpretaions where the values assigned to the variable in question varied. In examples involving quantification the value assignments for the narrow scope indefinites varied with values assigned to the variable bound by the quantifier.

It has been argued so far that the evaluation of noun phrases involves the following two possibly distinct issues relevant to scope: (i) the issue of the scope of the DC, and (ii) the issue of the scope of the variable.

3. The proposal

3.1. Evaluation indices

It is assumed here that the constituents of LF's are variables and predicative conditions placed on them. Truth conditions are stated in terms of functions in F assigning intensions to variables and functions in V assigning intensions to constants. The evaluation of both variables and predicative conditions varies relative to world parameters. Variables may be interpreted with respect to a single assignment function at varying worlds, as in the narrow scope indefinite cases discussed under modal quantification above, or they may be interpreted with respect to varying assignment functions at a single world, as in the case of narrow scope indefinites discussed under nominal quantification, or they may be interpreted with respect to a single assignment function at a single world, as in the case of wide scope indefinites discussed above. The parameters of variation for predicative conditions concern the world at which the extension of V for the predicate in question has to be checked. The world parameter of an evaluation function, as well as the assignment function evaluating a variable will be called evaluation parameters and will be indicated in LF's by evaluation indices. The value of an evaluation index specifies the value of the evaluation parameter of the expression it accompanies. Variation in scope results in variation in the value of evaluation indices. An expression e_1 counts as being in the scope of an expression e_2 iff the value of an evaluation index of e_1 is set relative to e_2. In (1a) the DC of the indefinite is within the scope of the intensional predicate iff the DC is to be evaluated at the world introduced by the matrix predicate. In the terminology used here, the world or modal index of the DC is set to the value introduced by the predicate. In (10), the indefinite has narrow scope with respect to the universal iff the assignment function parameter of the indefinite is dependent on the assignment function evaluating the universal.

The LF of a simple sentence such as (15a) is made up of the main predication (MP) (15c), given by the main predicative expression of the sentence, the variable contributed by the indefinite noun phrase (15a), and its DC (15d), given by the descriptive content of the noun phrase.

(15) (a) *A man left.*
(b) $x_1 \quad w$
(c) MP: $leave'(x_1)\ w$
(d) DC_1: $man'(x_1)\ w_m$

The w's here specify the values of the modal indices of the expressions they occur with. The modal index value on a variable, as in (15a) specifies the world whose domain the value for that variable must be chosen from. The modal index value on predicative conditions, such as those in (15c,d) specifies the world the predicate must hold of its arguments. The truth conditions of (15a) under the LF in (15b–d) are given in (16).

(16) (15a) is true at w wrt M iff there is an assignment function f such that
(i) $f(x_1)$ at $w \in V(leave')$ at w, and
(ii) $f(x_1)$ at $w \in V(man')$ at w_n.

A question that arises now is whether the modal index values of variables, their DCs and the MPs in which the variable is an argument may be set to values that are independent of one another. The discussion of the scope of the DC showed already that the modal index value of the DC of a variable is independent of the modal index value of the MP in which the variable occurs. The values of the indices of the variable and its DC are also independent both of one another and of the index of the MP. Relevant examples will be mentioned below.

Obviously, (15) when uttered in a neutral context is interpreted with all modal index values set to w_0, the world of the discourse. Alternative values for modal indices can be selected only if such values are made available in the discourse by being introduced by expressions such as modal operators, or intensional predicates, such as *believe*, *dream*, *want*, and intensional nouns, such as *belief* and *dream*.

3.2. Free and bound indices

Concentrating on the modal indices of main predications, I show now that they come in two varieties, *free* and *bound*. If an index is free, any previously introduced world is, in principle, an admissible value for that index.[2] If an index is bound, the value of that index must be set to a world introduced by some particular expression.

The modal indices of the MPs of main clauses are free; their default value is w_0, the world of the discourse, but this choice can be overridden in favor of previously introduced worlds. Consider the discourse in (17):

(17) *I had an unpleasant dream last night.*
The weather had turned unbearaly hot and my room was not air-conditioned.

Here the noun *dream* in the first sentence introduces a world w_d, a world in which the dreamt events occurred. The second sentence may be interpreted as being true of w_0, the world in which the discourse occurs, or as being true of w_d. I propose to account for these facts by having the world index of the MP of the second sentence be either w_0 or w_d. The second choice is possible only because w_d has been recently introduced. Under this indexing the second sentence is, in discourse structure terms, an elaboration of the first. The point of

[2] I am ignoring now certain constraints concerning (counter)factuality imposed by the morphology.

interest for present purposes is that the MP of a matrix clause may be indexed to a world other than w_0 in case such a world has been recently introduced and the discourse can be coherently interpreted as being about that world. Under this acount then, the structure of the interpretation of (17) under the two indexings is identical. What differs is the value of the modal index of the MP (and, most likely, the other elements) of the second sentence. A recently introduced world may serve as the value of a free index of an expression in subsequent discourse giving rise to discoursal modal subordination. Actual choices are, of course, limited by pragmatic and discourse coherence factors. Intensional predicates and nouns may have discourse scope because the worlds they introduce may serve as values to modal indices of expressions occurring in subsequent discourse. The second sentence in (17) may be interpreted as being 'about' the dream mentioned in the first sentence because the first sentence makes the 'dream world', w_d, available as a value for modal indices in subsequent discourse.

Unlike the modal index of MPs of matrix clauses, the modal index of the MPs of complements of intensional predicates or nouns as well as the modal indices of the MP in the modal quantificational structures are bound. Thus, the world index of the complement clause in (18) is necessarily set to w_J, the world introduced by the main predicate, as indicated informally in (18) by the modal indices subscripted to the clausal brackets.

(18) [John believes [that Mary is sick]$_{w_J}$]$_{w_0}$

The dependency of the modal index of complements on the world introduced by the expression the complement is an argument of is paralleled in the temporal realm by cases where the time reference of a complement is dependent on the time reference of the matrix.

In modal conditionals the Restrictor introduces a set of worlds, and the modal index of the MPs in the antecedent and in the conditional are bound to these worlds. (The fact that the LF in the Restrictor serves to identify the worlds in question, while the LF in the NS is interpreted as predicating something of them is the result of the way tripartite quantificational structures are interpreted.)

So far it has been claimed that predicative conditions as well as variables have modal indices whose values specify the worlds in which the predicate must hold of its arguments (in the case of predicative conditions), and the worlds from whose domain the variable must be given values (in the case of variables). The modal index of the MP of matrix clauses is free; its value can be set to w_0, the world in which the discourse occurs, or to some other previously introduced world. The modal index of the MP of arguments of intensional predicates as well as that of MPs in quantificational structures are bound. In the former case the modal index of the MP of the argument must be set to the world or worlds introduced by the intensional expression in question; in the latter case the value of the index is determined by properties of the immediate tripartite structure the MP is part of. The value of the modal index of the MP in the Nuclear Scope is determined on the basis of the interpretation of the Restrictor. Note that the value of a bound index is determined by properites of its local context, namely the predicate the MP is the argument of, and the domain over which one quantifies, in the case of MPs of Restrictor or Nuclear Scope expressions.

3.3. Scope of the DC

The facts concerning the scope of the DC are straightforwardly accounted for under the assumption that the evaluation indices of DCs are free. The freedom of the modal indices of the DC of noun phrases parallels the freedom of their temporal index argued for in Enç (1986).[3] Under present assumptions this amounts to claiming that the modal index of the DC of a noun phrase can be set, at least in principle, to any world the context makes available. In particular, the value of this index may be the same as that of the index of the MP or it may be some other world that is salient in the context. (This is exactly the same as the behavior of temporal indices in Enç's analysis.) Note that under this approach the DC of a noun phrase e_1 has 'narrow scope' with respect to an intensional predicate or modal e_2 iff the modal index of the DC is given as value the worlds introduced by e_2. The DC has 'wide scope' with respect to a modal or an intensional predicate iff the value of its modal index is not set to the worlds introduced by them. The two readings of (1) would have their indices set to the values given in (19) and (20):

[3] There may well be noun phrase types whose DCs are bound to the parameters of their main predications. These would be noun phrases such as bare plurals, that have necessarily narrow scope with respect to expressions in higher clauses.

(19) [John believes [that [a friend of mine]$_{w_J}$ is a crook]$_{w_J}$]$_{w_0}$
(narrow scope)

(20) [John believes [that [a friend of mine]$_{w_0}$ is a crook]$_{w_J}$]$_{w_0}$
(wide scope)

The choice of w_J in (19) is available because the DP occurs within the complement of *believe*, which introduces w_J; the choice of w_0 in (20) is available because the world of the discourse needs no special introduction.

Under this view, the two readings have representations that are structurally identical; they differ in the values of the evaluation indices on the DC, and not in the structural position of the DC relative to the predicate *believe* at either the syntactic or the semantic level. The fact that the DC occurs within the argument of *believe* makes the choice in (20) possible but not necessary.

3.4. Scope of the variable

Recall that the question of the scope of the variable as defined here concerns the possible dependency between assignment functions for one variable x on varying assignments for another variable y. In case of variable dependency x is evaluated with respect to assignment functions that co-vary with the assignment functions evaluating y. Thus, in (10) the two interpretations result from the option of evaluating the variable contributed by the indefinite by the base assignment function f or by the set of functions f_{NS} which extend the functions in the set f_R. The former choice gives the wide scope reading and the latter the narrow scope reading.

I propose that the function that evaluates a variable is one of its evaluation indices, and is indicated in LF on a par with modal indices. The choices are determined by the position and role of the variable in LF, which in turn are dependent on surface structure considerations. The function index of a variable contributed by a quantificational DP is bound: it has to be the set of functions f_R. The function index of non-quantificational variables is free. For a variable contributed by an indefinite in the NS the available options are (i) the base function f, which, just like the world of discourse, is always an option for a free function index, and (ii) the functions in the set f_{NS}. This latter option is available because the variable in question is in the NS in a nominal quantificational structure. The two options are represented as in (21a) and (21b):

(21) (a)

Restrictor:	$\forall x_1$	Nuclear Scope:
x_1 f_R		x_2 f
DC_1: *student* $'(x_1)$		DC_2: *I-E language*$'(x_2)$
		MP: *speak* $'(x_1, x_2)$

Restrictor:	$\forall x_1$	Nuclear Scope:
x_1 f_R		x_2 f_{NS}
DC_1: *student* $'(x_1)$		DC_2: *I-E language*$'(x_2)$
MP: *speak* $'(x_1, x_2)$		

The effect of having the value of the function index of x_2 be f in (21a) is the same as having x_2 be bound by text-level existential closure; the effect of having the value of the function index set to f_{NS} is the same as having the variable be bound by NS-level existential closure. The truth conditions of (21a) and (21b) are as in (22):

(22) (a) The LF in (21a) is true in w wrt M iff there is an assignment function f such that every assignment function f_R that extends f such that $f_R(x_1) \in V(student')$ has an extension f_{NS} such that $f(x_2) \in V(I\text{-}E\ language')$ and $< f_{NS}(x_1), f(x_2) > \in V(speak')$.

(b) The LF in (21b) is true in w wrt M iff there is an assignment function f such that every assignment function f_R that extends f such that $f_R(x_1) \in V(student')$ has an extension f_{NS} such that $f_{NS}(x_2) \in V(I\text{-}E\ language')$ and $< f_{NS}(x_1), f_{NS}(x_2) > \in V(speak')$.

It is crucial to note that the representations of the two readings are structurally identical; the only difference concerns the value of the function index on x_2. In structural approaches the choice of evaluation function for a variable is unambiguously determined by the structural position of the DP contributing the variable at LF or the structural position of the variable at the level of DRS. The same is true for the value of the modal parameter. In the approach proposed here the value of these parameters is encoded directly in the semantic representation in the form of values for evaluation indices. Structural considerations are relevant only in determining what options are available.

In examples involving modal quantification the cases introduced by the Restrictor are worlds which must serve as values to the modal index of the MPs of both Restrictor and NS. The modal index of variables contributed by indefinites in either the Restictor or NS is free and therefore these worlds may or may not serve as values of the modal index of variables in the Restrictor or the NS. If they do, the assignment function giving values to that variable will range over the worlds introduced by the Restrictor resulting in narrow scope readings. If they do not, the variable will be rigid with respect to these worlds resulting in wide scope readings.

In the approach proposed here, wide and narrow scope readings for both the DC and the variable are distinguished by evaluation index values and not by differences in the hierarchical position of the relevant expressions at LF or DRT. Consequently, neither structural level needs to be tampered with (by QR in the case of LF, and by structural accommodation in the case of DRS) in order to account for scope facts.[4] The consequences of this approach regarding limits on the scope domains of indefinites and *every* are dealt with in the next section.

4. Scope of *every* vs. scope of indefinites

I will be concerned here with the scope domains of indefinites and the universal determiner *every*. I first review the empirical generalizations established in the literature and then show how they are accounted for under the present proposal. (For relevant discussion, see Fodor–Sag 1982; Farkas 1981; Ludlow–Neale 1991; Abusch 1994.)

Based on the fact that (23a) has no reading in which officials co-vary with committee members, while (23b) has a reading in which a single high placed official is involved,

(23) (a) *A high placed official claimed that John talked to every member of the committee.*

(b) *Every member of the committee claimed that John talked to a high placed official.*

the following generalizations emerge:

A. The scope of *every* and modal quantifiers is upward clause-bounded with respect to indefinites or other quantifiers.

B. The scope of indefinites is upward unlimited.

[4] For different proposals dispensing with (most) instances of QR see Beghelli (1993) and Dobrovie-Sorin (1993).

The fact that modals pattern like *every* is shown by the interpretations of the examples in (24): the modal may not have scope over the variable contributed by the indefinite in (24a), *every* may not have scope over the modal in (24b), but the variable contributed by the indefinite may have scope over the modal in (24c).

(24) (a) *A student thinks that John must leave.*
(b) *It is possible that every candidate wins.*
(c) *It is possible that a candidate wins.*

In the present approach, the generalization in A reduces to the claim that the cases introduced by the Restrictor in a quantificational structure are not possible index values for variables contributed by expressions occurring in clauses that structurally command the quantificational expression. The generalization in B reduces to the claim that a variable in the NS need not inherit the index of the cases introduced by the Restrictor.

Note now that the latter generalization has already been captured by the assumption that the function index of a variable contributed by an indefinite is free. The value of such an index may be the base function, giving the wide scope reading, or the set of functions f_{NS}, giving the narrow scope reading. The availability of the wide scope reading is ensured by the availability of the base function as an index value for indefinites. The behavior of values for the function index parallels that of the modal and temporal index of the DC, which can be set to any contextually salient value. The possibility of widest scope for both the variable contributed by an indefinite and its DC are the result of the base function and the world of the discourse being salient potential values for free indices.

The main advantage of this approach over an LF-based approach is that the unlimited upward scope of DCs and of variables contributed by indefinites is accounted for without having to posit unlimited QR (or NP-preposing), and without having to assume long distance binding by text-level existential closure as in Abusch (1994).

The generalization in B can be captured by assuming the constraint in (25):

(25) A value v is accessible as a possible value to an index of an expression e_1 iff v as been introduced prior to e_1

Let us assume now that at the discourse level priority is defined in terms of temporal sequencing. This allows any index value v contributed by an expression e_2 to be accessible to expressions in subsequent discourse. This assumption accounts for (17), as well as the other cases of modal subordination discussed in Farkas (1993). Constraints concerning discourse coherence and larger discourse structure will play a role in accounting for the various limits on the discoursal scope of various expressions.

At the level of complex sentences, let us assume that the constituents of a clause S_1 are prior to the constituents of a clause S_2 iff S_1 c-commands S_2 in surface structure. Under these assumptions, the generalization in B is accounted for.

Finally, let us assume that clause-mates are simultaneous, i.e., that the index value introduced by an expression e_2 is accessible as a value to any expression e_1 that occurs in the same minimal clause as e_2. This predicts the possibility of 'inverse' scope in (26).[5]

(26) *A proofreader read every paper.*

Under these assumptions priority at the sentence level corresponds to c-command defined on clausal domains in surface structure.

A welcome consequence of this approach is that it accounts straightforwardly for what turns out to be a paradox in structural accounts of scope, namely the fact that the DC of a distributive noun phrase is upward unlimited, while the scope of the distributive quantifier contributed by the noun phrase is upward clause-bounded. (See Farkas 1993 and Ludlow–Neale 1991 for discussion.) Consider (27):

[5] There are cases when command seems to matter even within a minimal clause. Thus (i) may not be interpreted with the indefinite within the scope of the distributive while (ii) can.

(i) *John showed somebody/a student every picture.*
(ii) *John showed every picture to somebody/a student.*

É. Kiss (pc) has also noted that the relative scope of clause-mate quantificatifiers is fixed by surface c-command as well. In the face of such facts one could either adopt a strict c-command based priority relation within the clause as well and relax it only for the cases where inverse scope within a clause obtains or one can adopt the permissive version suggested in the text and restrict it in the cases where inverse scope within a clause does not obtain. The former alternative owes an explanation for the instances where inverse scope is possible; the latter owes an explanation for those where it is not.

(27) *Mary thought that a witch claimed that every person in this room had contact with her.*

Here *a witch* cannot co-vary with assignments given to the variable contributed by the distributive noun phrase. This is predicted by the present theory since the index contributed by the distributive is not accessible to the indefinite. The DC of the distributive, on the other hand, may have scope over both *claim* and *think*. This, again, is predicted in the present proposal, since the scope of DCs is upward unlimited.

This example illustrates the independence of the scope of the DC from the scope of the quantifier. The independence of the scope of the DC from the scope of the variable in the case of indefinites is illustrated in (28).

(28) *I would be happy if somebody who is actually rich would have been poor.*

(Similar cases are discussed in Abusch 1994 as well as in Farkas 1985.) This example has a reading where the variable contributed by the indefinite is indexed to the worlds introduced by the Restrictor, thus allowing its value to vary from world to world, but the modal index of the DC of the indefinite must be set to w_0. Here then the variable has narrow scope with respect to the worlds introduced by the Restrictor, while its DC has wide scope with respect to them.[6]

The main advantage of the proposed approach to scope is that it accounts straightforwardly for both the upwardly unlimited scope of DCs and indefinites both within a sentence and beyond it without having to resort to unbounded movement or unbounded binding. The unlimited scope of these expressions is the result of the availability of index values introduced in prior discourse. The selection of values for the evaluation indices of these expressions is reminiscent of antecedent selection for non-reflexive pronouns. In the

[6] There is a complication with examples such as

(i) *John thinks that every friend of mine cheated.*

which cannot be interpreted with the quantifier having wide scope and the description having narrow scope with respect to *think*: the sentence cannot be understood as quantifying over all and only those actual individuals who are friends of mine according to John, excluding those individuals who exist only in w_j and are friends of mine there. It may be that there is a constraint requiring the DC and its quantifier to share their modal index so that one does not have to look at several worlds when determining what constitutes a case for quantification.

case of expressions whose evaluation indices are bound, i.e., expressions whose index value is predetermined, there is a local relation between the element introducing the value and the constituent whose index must be set to the value in question and the binding of the index is intimately connected to the interpretation of the construction the expression with the bound index is an immediate constituent of.

In approaches where scope is determined by hierarchical relations at LF the upward unlimited scope of DCs and indefinites can be accounted for either by allowing unlimited QR (or NP-preposing) or by allowing existential quantifiers (but not universal ones) to bind a variable indefinitely lower, as proposed in Abusch (1994). Both alternatives involve an unbounded relation of a type that has no precedent elsewhere: either an unconstrained movement or an unconstrained binding relation between quantifier and variable.

The fact that the scope of the quantifier *every* is upward clause-bounded follows in the present approach from two assumptions that appear relatively innocuous, namely the assumption that *every* is not raised, or at least not raised beyond its clause, and the assumption that index values are not available prior to their introduction. Problematic data for these assumptions are discussed below.

It has been noted in the literature that *each* and less easily, *every*, may take scope over an indefinite in a higher clause.[7] Thus, at least for some speakers, the indefinite in the matrix clause in (29) may receive a non-rigid interpretation with respect to the cases identified by the Restrictor in the subordinate clause, namely the relevant invited speakers.

(29) *A/some student made sure that each/every invited speaker had a ride.*

It is not clear to me yet what conditions allow or facilitate this liberal scope. At present I can only make some empirical observations. First, while the scope of these distributives may cross one sentence boundary, it may not cross more than one. Thus, the matrix subject in (30) cannot be dependent on the cases introduced by the distributive noun phrase:

(30) *A/some student made sure that John arranged that each/every invited speaker has a ride.*

[7] Thanks to Ivan Sag for reminding me of these cases. See Ioup (1977) for relevant discussion.

Second, it appears that the nature of the matrix predicate influences the possibility of the distributive to take extra wide scope. Thus, speakers who accept the relevant reading of (29) cannot interpret (31) with the distributive having extra wide scope:

(31) *A/some journalist or other said that each/every candidate won.*

Finally, *each* seems to take extra wide scope significantly more easily than *every*. Other quantifiers such as *all*, *several*, *most*, or cardinal numerals do not seem to allow for this possibility.

The extra wide reading of the distributive in (29) is clearly problematic for the present approach, under the assumption that higher clauses are prior relative to lower clauses. Given the empirical observations above, however, one should certainly not abandon the fundamental observation formulated in generalizations A and B above, i.e., one should not conclude that the scope of distributives, just like that of indefinites, is upward unbounded. I suspect that the correct solution will make crucial use of the emphatically distributive character of *each* and *every* and cannot be reached without a deeper understanding of the nature of distributivity and its interaction with quantification.[8] Progress in this area will also shed light, I think, on the scopal properties of various types of plural noun phrases as well as of determiner types that have not been included in the present discussion, such as cardinals and amount denoting determiners. My main goal here was to show that an indexical treatment of both DC and variable scope successfully handles some essential properties at the level of the discourse and the complex clause, and therefore that it is worth pursuing with respect to the many questions that remain open.

[8] Another possibility would be to capitalize on the presupposed nature of the domain of these two quantifiers (discussed in Moltmann 1994). This, however, is less promising since it does not distinguish *each* and *every* from *most*.

References

Abusch, D. 1994. The scope of indefinites. In: Natural Language Semantics. Vol. 2, No. 2, 83–135.

Beghelli, F. 1993. A minimalist approach to quantifier scope. In: NELS 23.

Carlson, G.N. 1987. Same and different: some consequences for syntax and semantics. In: Linguistics and Philosophy 10: 531–65.

Dobrovie-Sorin, C. 1993. On the denotation and scope of indefinites. Ms. CNRS, Paris.

Enç, M. 1986. Toward a referential analysis of temporal expressions. In: Linguistics and Philosophy 9: 405–26.

Farkas, D.F. 1981. Quantifier scope and syntactic islands. In: CLS 17: 59–66.

Farkas, D.F. 1985. Intensional Descriptions and the Romance Subjunctive Mood. Garland, New York.

Farkas, D.F. 1993. Modal anchoring and NP scope. Linguistics Research Center working paper. LRC-93-08. Santa Cruz, CA.

Fodor, J.D.–Sag, I. 1982. Referential and quantificational indefinites. In: Linguistics and Philosophy 5: 355–98.

Heim, I. 1982. The semantics of definite and indefinite noun phrases. Doctoral dissertation, MIT.

Ioup, G. 1977. Specificity and the interpretation of quantifiers. In: Linguistics and Philosophy 1: 233–45.

Kamp, H.–Reyle, U. 1993. From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory. Kluwer Academic Publishers, Dordrecht, Holland.

Kratzer, A. 1977. What "must" and "can" must and can mean. In: Linguistics and Philosophy 1: 337–55.

Kratzer, A. 1981. The notional category of modality. In: Eikmeyer, H.-J.–Rieser, H. (eds): Words, Worlds, and Contexts, 38–74. Walter de Gruyter, Berlin.

Ludlow, P.–Neale, S. 1991. Indefinite descriptions: in defense of Russell. In: Linguistics and Philosophy 14: 171–202.

Moltmann, F. 1994. Intensional verbs and quantifiers. Ms. UCLA.

Poesio, M.–Zucchi, A. 1992. On telescoping. In: Barker, C.–Dowty, D. (eds): SALT II, 347–67.

Roberts, C. 1989. Modal subordination and pronominal anaphora in discourse. In: Linguistics and Philosophy 12: 683–721.

Address of the author: Donka F. Farkas
Stevenson College
UCSC
Santa Cruz, CA 95060
USA

Acta Linguistica Hungarica, Vol. 42 (3–4), pp. 305–323 (1994)

LOGICAL SEMANTICS OF DEFERRED INFORMATION*

LÁSZLÓ PÓLOS

1. Introduction to deferred information

The notion of deferred information is best introduced in connection with presuppositions. In what follows we briefly recapitulate some of the relevant facts that has been established (Kálmán–Pólos–Szabó 1989; Kálmán–Pólos–Szabó 1990; Kálmán–Szabó 1990; Kálmán 1990) about presuppositions and deferred information.

1.1. Presuppositions

In the Tarskian–Montagovian tradition of formal semantics the meaning of declarative sentences is characterized in terms of their truth conditions. In this tradition any new (true) sentence we learn constrains the possibilities of what the real world could look like. We can identify the proposition expressed by a declarative sentence (i.e., the meaning of the sentence) with the set of possible worlds (with those worlds where the proposition is true).

Some declarative sentences express propositions only conditionally.

(1) *The present king of France is bald.*

is a declarative sentence, but expresses a proposition only if there exists a (unique) king of France. If the world is such that France lacks any king whatsoever, the sentence does not express a proposition, i.e., it is neither true nor false. The existence of a (unique) king is the presupposition of the sentence.

A tentative, preliminary definition of the presupposition of a declarative sentence is the following: The presupposition is a proposition that has to be true, otherwise the sentence does not express a proposition, i.e., it can be neither true nor false.

* The research presented in this paper was supported by the grant No. 3217/94 of the Hungarian Ministry of Education and the Dutch National Science Foundation (NWO Project No. PG550,334)

1.2. Presupposition accommodation

The definition above is over-simplified, and has certain shortcomings. We illustrate this by the following: The truth of the presupposition depends on *one* particular world, namely the actual one. Suppose that we have two declarative sentences ϕ_1 and ϕ_2 and suppose furthermore that both of these sentences have presuppositions, p_1 and p_2 respectively. If in the actual world both p_1 and p_2 happen to be false, then neither of ϕ_1 and ϕ_2 express a proposition, neither of them can be either true or false, so the truth-conditions and hence the meanings of ϕ_1 and ϕ_2 coincide. But we are somewhat reluctant to accept that the sentence:

(1) *The present king of France is bald.*

means the same as the sentence:

(2) *The vice-president of the United Kingdom smokes cigars.*

A bit more realistic view can be developed if we opt for a semantics where truth and falsity of propositions in general and presuppositions in particular depend on a set of possible worlds, sometimes called information states. In such a setup a proposition is true in an information state iff all elements of the information state are elements of the proposition too. If we define falsity in a dual manner, a proposition is false in an information state if no element of the information state is an element of the proposition. Truth and falsity both are relativized to information states, and we can preserve the principle *tertium non datur* only if we admit that some formulae no longer express propositions. They fail to do so if whatever they express is neither true nor false.

A declarative sentence is accepted (i.e., expresses a proposition) in an information state if all the presuppositions of the sentence are true in the information state. Such a semantic setup guarantees that if two sentences have different presuppositions, their truth-conditions, as well as their meanings, are different.

The following example shows that even this setup is not realistic enough. Suppose for example that we see Peter enter the room in his pullover. One can well start a discourse with the sentence:

(3) *Peter left his overcoat in the hall.*

It would be quite meaningful. The presupposition of this possessive construction is the proposition expressed by the following sentence:

(4) *Peter has an overcoat.*

But this presupposition is not made true by the initial (empty) information state at all. As a matter of fact we just learned form the sentence

(3) *Peter left his overcoat in the hall.*

that he has got one.

What is going on in this case? It is not the preceding information state that makes the presupposition true, but the one that follows the incorporation (see Kálmán 1990) of the sentence. The moral of this simple example is: if the presupposition is not made true by the previous context the interpreter is willing to **accommodate** the required pieces of information.

1.3. How does deferred information work?

The possibility to accommodate presuppositions, if needed, threatens to empty the notion of presupposition. (We started with the strong requirement that the presupposition **has to** be true and ended with a much weaker position: if the presupposition is not true, take a similar information state which looks like the one in question but makes the presuppositions true.) But fortunately presupposition accommodation is not always possible.

A presupposition can be true in an information state if all elements of the information state are in the presupposed proposition, can also be false if no element of the information state is in the presupposed proposition, and there is a third possibility: neither true nor false, i.e., some elements of the information state are in the presupposed (would-be) proposition and some are not. In the first case there is no need for presupposition accommodation, in the second case there is no possibility to accommodate. Only the third case provides both need and room for presupposition accommodation. We need to accommodate, to make the sentence express a proposition, and we can accommodate by throwing away all elements of the information state that are not elements of the presupposed proposition.

Presupposition accommodation is sensitive to the context: some pieces of information are easily acccommodable in some contexts whereas some other pieces are not at all accommodable. Only those pieces of information are accommodable into an information state that are somehow **already** present in it.

If not in the information state then in its *aura*. We follow Kálmán (1990) in terminology and call these pieces **licensed** by that context. To illustrate this point take the following (already somewhat standard) example:

(5) Peter got married yesterday. The priest spoke nicely.

(6) Peter got married yesterday. The dog barked loudly.

The priest is a definite noun phrase and it presupposes that there is a unique priest in the context. But the previous context does not provide any priest. Still the existence of a unique priest can be accommodated into the context of a marriage. Among other, more secular alternatives, Peter might have married in church. If he did so, a unique priest entered the context, namely the one who celebrated the ceremony. The presence of a unique priest is a **deferred piece of information** in the information state we get after processing the first sentence of the relevant discourse. That again illustrates the point that presuppositions can be accommodated, if they are (as in our case) not implied by the previous context, and not excluded by that context either.

The dog just like *The priest* carries a presupposition. The presupposition is that there is a unique dog in the context. And the previous context does not provide any dog. But the context does not exclude the possibility of a dog being present, either. Still the existence of a (unique) dog cannot be accommodated. The context does not licence such an accommodation. The reason is that dogs (as well as any particular dog) are absent from the context. This piece of information is not deferred.

1.4. Side effects of the use of deferred information

We argued above that it is the stock of deferred information in an information state that determines whether a particular piece of information can or cannot be accommodated to the information state. This is—so to say—the effect that indicates the presence of deferred pieces of information. In presupposition accommodation deferred information exhibits itself not only via this effect but also via certain side effects. When a presupposition accommodation takes place, new pieces of information may become available, which were not implied by the previous context, nor by the accommodated information and not even by the conjunction of the previous context and the presupposition.

The accommodation of the existence of a (unique) priest relied (among other things) on the possibility that people might marry in church. Neither of the sentences in the discourse nor their conjunction implies that Peter did marry in church, but as a side effect of presupposition accommodation the interpreter tends to believe that in fact that is what happened.

To explain the side effects of presupposition accommodation one has to assume that several deferred pieces of information are **lumped together**, and all the content of such a lump becomes available if one of the pieces in a lump are used as a licence of a presupposition, i.e., is accommodated.

2. SitUp Setup

In the introduction above we sketched the notion of information state already (following Veltman 1985, or Veltman 1995 for example). Even though that notion of information state is clear enough, for our present purposes it is not satisfactory. We need a systematic way to represent presuppositions in information states in order to picture the effects and side-effects of presupposition accommodation. Furthermore information states should also be able to picture the *aura*, i.e., the lumps of lexical information.

Situation theory (Barwise–Cooper 1994) offers the concept of restricted objects. The restriction is a proposition, and the truth of that proposition is the *conditio sine qua non* of the non-degenerate interpretation of the restricted object. If the restricted object is the representation of the meaning of a (declarative) sentence and the representation has a non-degenerate interpretation if and only the representation is a proposition, then the restriction has the very same semantic effect as presuppositions. In what follows we exploit this similarity and use situation theory to define a new notion of information states.

Situations of Barwise–Cooper are not perfectly suited to formalize information states: we have to modify the notion of situation in order to be able to represent disjunctive pieces of information. Furthermore to interpret restrictions as presuppositions requires a substantial number of modifications. Roughly speaking, restrictions are considered as presuppositions, i.e., propositions that have to be made true by the (previous) context. Accordingly, not only extensions but also anti-extensions are associated to types in all situations. A further consequence of this deviation that proposition-hood has to be made situation dependent too. Since the truth or falsity of the restrictions depends on the situation, so does proposition-hood.

2.1. Key concepts of situation theory

Both formal and natural languages are composed of sentences. These are the minimal strings of the language that express content and are suited to assume a truth value. One cannot add less **content** to a situation than the content associated with an atomic sentence. The **representation** of atomic sentences is called **infon.** Consider the atomic sentence

(7) *Romeo loves Julia.*

This sentence is composed of a two-place predicate and two objects filling the argument slots of the predicate. We could represent the sentence by the sequence

$$(Loves, Romeo, Julia)$$

but as a representation, the sequence should appear as an object in its own right, different from the string

(8) (Loves, Romeo, Julia).

To avoid confusion, the denotation of the representation is identified by double angle brackets:

$$<< Loves, Romeo, Julia >>$$

Information is not only contingent on truth; absent information may represent lack of knowledge, rather than lack of truth, so we need a partial framework. To express that a piece of information is true, we assign a positive truth value (= 1, or **polarity** in situation-speak) to it; if it is false, the polarity is negative (= 0). Lack of knowledge is represented by the absence of information, and hence by the absence of the corresponding infons in the situation. For example, the fact that we have not learned whether Romeo loves Julia does not imply that Romeo does not love Julia. Accordingly, the content of the sentence *Romeo loves Julia* is represented by:

$$<< Loves, Romeo, Julia; 1 >>$$

and the content of *Romeo does not love Julia* is represented by an **infon**:

$$<< Loves, Romeo, Julia; 0 >>$$

2.1.1. Situations and infons

From a **relation** and an appropriate number of **objects** and a **polarity** we can build an **infon**. Infons play an important role in the representation of sentence meaning. A situation is characterized by the infons it supports. To express that in a particular situation (say in the play of Shakespeare, σ for short) this infon holds, we use the support relation:

$$\sigma \models << Loves, Romeo, Julia; 1 >>$$

2.1.2. Propositions

In situation theory we call a string like the last one a proposition. A situation either supports an infon or it does not support it, so these sort of propositions are either true (if the support relation holds) or false (if it does not hold). If ϕ is a sentence of the language we want to represent, and σ is a situation, then $\sigma[\![\phi]\!]$ denotes the situation that we get by updating σ with the information associated with ϕ.

2.1.3. Types and assignments

An alternative way to look at the proposition

$$\sigma \models << Loves, Romeo, Julia; 1 >>$$

is that it classifies the play (i.e., the situation σ) as one of those that support the infon:

$$<< Loves, Romeo, Julia; 1 >>$$

Generally we can consider propositions as classifications: For this reason we introduce a new notation and write $\sigma : \phi$ to denote that σ is of the ϕ-supporting type of information states. $\sigma : \phi$ is a proposition, involving types of situations. We get these types by λ-abstraction: if the type ϕ is a proposition from which a situation is λ-abstracted, and : is interpreted as function application, then we have

$$(s : \lambda s[p]) = p$$

We generalize the notion of types beyond the case discussed above because this will give us a tool for classifying arbitrary objects. The information that a particular object belongs to an information type would expresses a generalized proposition, as it were.

Types result from lambda abstraction over propositions. Such propositions may contain more than one variable. One might need to λ-abstract some or all of these variables. The term **assignment** is a shortcut for variable assignments to n-tuples of objects. With such assignments, one can simulate iterated lambda-abstraction by considering alternative assignments.

2.1.4. Propositions β

Assignments can be classified by information type. Now, if α is an assignment and τ is an information type, then $\alpha : \tau$ may or may not be a proposition, depending on whether the assignment is appropriate for the type or not. This uncertainty gives rise to considerations of appropriateness, and motivates the introduction of another concept: preproposition.

2.1.5. Prepropositions

Prepropositions, loosely speaking, are possibly inappropriate, pseudo-propositions. Not all syntactically well-formed expressions of natural languages can carry meaning; consequently, they can be neither true nor false. Prepropositions have the form of propositions, but they need not have a truth value.

2.1.6. Truth

The general form of a proposition is $\alpha : \tau$ where α is an assignment and τ is a type. We can specify truth conditions for these propositions in terms of extensions of types. Unlike propositions of the form $\sigma \models \phi$, general propositions are not always true or false, since their truth conditions might depend on specific information states.

To handle this complication, we introduce two additional concepts, namely the notions **extension** and **anti-extension** of types. Both extension and anti-extension are sets of assignments.

The general strategy of situation theory is to provide a second order, type free framework. Due to this strategy the notion of truth as well as falsity is internalized in term of the types TRUE and FALSE.

3. Updates

To bring the static, denotational tradition of situation semantics and the dynamic tradition together our strategy is the following:

– First an update operator $\oplus$ is introduced. This operator takes a situation σ and an infon ϕ and creates a new situation $\sigma \oplus \phi$. The denotation of several declarative sentences are infons, and the most straightforward manner to model their dynamic semantics is to update the actual situation with their denotations.

– Secondly postulate the existence of an empty situation $\oslash$, i.e., a situation that does not support any infon whatsoever. (As an information state this object represents the information state of complete ignorance.)

To implement the changes described in the first and the second step, the following new postulates are required:

(A_1) $\oslash \in SIT$

(A_2) $\oslash \not\models \xi$ for all $\xi \in INFON$

(A_3) $\sigma \in SIT$ and $\xi \in INFON$, then $\sigma \oplus \xi \in SIT$

(A_4) for all $\xi, \sigma, \eta, \xi \in INFON$ and $\eta \in INFON$ and $\sigma \in SIT, \sigma \models \xi$ implies that $\sigma \oplus \eta \models \xi$ and $\sigma \oplus \eta \models \eta$ and for all$\sigma' \in SIT$ if $\sigma \models \xi$ implies that $\sigma' \models \xi$ and $\sigma' \models \eta$ then $\sigma \oplus \xi$ is a part of σ'

where the part-of relation is defined as usual:

$$\sigma \text{ is a part of } \sigma' \text{ iff for all } \xi\ \sigma' \models \xi \text{ implies that } \sigma \models \xi$$

– The third step is to specify the update conditions operating on situations. Our main concerns are: 1) updates with disjunctive pieces of information; 2) the presence of lexical information.

With the formal machinery in place, we are now ready to discuss its use in representing deferred information. Kálmán (1990) introduces three sources of deferment: disjunctivity, lexical information and genericity. In our present paper we deal with the first two and it is left to a subsequent paper to study the logical semantics of deferred information related to generic sentences. We start with the most obvious source of deferred information: the lexicon. For this reason we show how lexical information enters the representation.

3.1. The lexical information

The update of a situation does require the import of lexical information. In our framework the set of lexical items is composed of three (not necessarily disjoint) parts: primitive relations, atoms, and primitive types.

Let LEX denote the operation that maps all the situation theoretic objects into the power set of lexical items. The set of lexical items of situation theoretic objects is defined recursively. The recursion follows the axioms that we introduced (see Pólos–Masuch 1995) to characterize these objects:

$$LEX : OBJ \longrightarrow \mathcal{P}(\Lambda)$$

The operation is to be defined by the following conditions.

1. if $r \in R_0$, then $LEX(r) = \{r\}$
2. if r is a relation of the form $\lambda x, \pi[\xi]$, then $LEX(r) = LEX(\xi)$
3. if r is a relation of the form $r'|p$, then $LEX(r) = LEX(r')$
4. if ξ is an infon of the form $<< r, a_1, \ldots, a_n; \pi >>$, then
$$LEX(\xi) = LEX(r) \bigcup LEX(a_1) \bigcup \ldots \bigcup LEX(a_n)$$
5. if ξ is an infon of the form $\neg\xi_1$, then $LEX(\xi) = LEX(\xi_1)$
6. if ξ is an infon of the form $\xi_1 \wedge \xi_2$, then
$$LEX(\xi) = LEX(\xi_1) \bigcup LEX(\xi_2)$$
7. if ξ is an infon of the form $\xi_1 \vee \xi_2$, then
$$LEX(\xi) = LEX(\xi_1) \bigcup LEX(\xi_2)$$
8. if ξ is an infon of the form $\xi_1|p$, then $LEX(\xi) = LEX(\xi_1)$
9. $LEX(\oslash) = \emptyset$
10. $LEX(\sigma \oplus \xi) = LEX(\sigma) \bigcup LEX(\xi)$
11. if σ is a situation, and p is a proposition, then
$$LEX(\sigma|p) = LEX(\sigma)$$
12. if p is a proposition of the form $\sigma \models \xi$, then $LEX(p) = LEX(\sigma)$
13. if f is an assignment, then
$$LEX(f) = \{a \in \Lambda \mid a \in LEX(b) \; for \text{ some } \; b \in \mathcal{R}(f)\}$$

where $\mathcal{R}(f)$ is the range of the assignment f.

14. if p is a proposition of the form $f : \tau$, then
$$LEX(p) = LEX(f) \bigcup LEX(\tau)$$
15. if both p and q are propositions, then $LEX(q|p) = LEX(q)$
16. if p is a proposition of the form $\neg q$, then $LEX(p) = LEX(q)$
17. if p is a proposition of the form $\sim q$, then $LEX(p) = LEX(q)$
18. if p is a proposition of the form $r \wedge q$, then
$$LEX(p) = LEX(r) \bigcup LEX(q)$$
19. if p is a proposition of the form $r\&q$, then
$$LEX(p) = LEX(r) \bigcup LEX(q)$$
20. if p is a proposition of the form $r\ V\ q$, then
$$LEX(p) = LEX(r) \bigcup LEX(q)$$
21. if p is a proposition of the form $r\mathrm{v}q$, then
$$LEX(p) = LEX(r) \bigcup LEX(q)$$
22. if $\tau \in TYPES_0$, then $LEX(\tau) = \{\tau\}$
23. $LEX(\lambda x[\xi]) = LEX(\xi)$
24. if $\tau \in TYPES$, then $LEX(\tau|p) = LEX(p)$

The lexical content of a lexical item is given in the form of a situation. From now on we write $LCO(\alpha)$ for the situation expressing the content of the lexical item α. The lexical contents do not in general contribute to a situation directly. But they are present as constraints. A situation σ *should* be compatible with all information states of the form $LCO(\alpha)$ if $\alpha \in LEX(\sigma)$. (This is the formal meaning of what we called *aura* above.)

We implement this constraint formally in terms of restrictions. A lexical item α is to be substituted by an item restricted by the proposition that expresses this lexical constraint:

$$\alpha \mid \alpha \in LEX(\sigma) \longrightarrow \sigma \oplus \bigwedge LCO(\alpha) \neq \otimes$$

Note that the LEX operation was defined such that restrictions do not contribute to the lexical content, and accordingly we need not apply the above substitution repeatedly. Note furthermore that the implementation of the lexical constraint will not tell us exactly which specific situation express the lexical contents of certain lexical items. This remains an empirical question for linguists to answer.

3.2. Updating situations with complex infons

There are four different infon constructions to be considered here: negation of atomic infons, conjunction and disjunction of arbitrary infons, and restricted infons. In the case of an atomic infon, the receiving agent adds the infon to its information state, and adjusts the state of the lexicon so that it contains the new items.

3.2.1. Negation

Negation is defined only for atomic infons, and involves a change in the infon's polarity. If ξ is an atomic infon, then ξ is of the form

$$<< r, a_1, \ldots, a_n; \pi >>$$

and the negation of ξ is

$$<< r, a_1, \ldots, a_n; 1\text{-}\pi >>$$

If the situation σ is updated with a negated infon $\neg\xi$, then either σ supports ξ already, in which case the update yields a contradictory, and hence absurd, situation, or σ did not support ξ, in which case the updated information state must support $\neg\xi$ plus all structural information associated with $\neg\xi$.

3.2.2. Conjunction

One main argument in favour of a situational representation relies on the fact that sequencing and conjunction work out differently when information flows. Let $\bigwedge \Phi$ and $\bigwedge H$ be two conjunctions of infons associated with ϕ and η respectively. Updating a situation σ with a text ϕ means to add the infons expressing the information associated with ϕ to σ. In case of a conjunction, we have:

$$\sigma \oplus (\bigwedge \Phi \wedge \bigwedge H)$$

and in case of a sequence, we have:

$$\sigma \oplus \bigwedge \Phi \oplus \bigwedge H$$

In the first case there is only one update step, and all the expressions that refer back to preceding situation; in the second case, the expressions in ϕ refer to σ, but the expression in H refer to $\sigma \oplus \bigwedge \Phi$.

3.2.3. Disjunction

In case of a conjunctive update, all conjuncts (plus the associated, structural information) is added to the situation. In case of a disjunctive update, the situation looks more complicated, since it is not obvious which infons are to be added. Consider the following example:

(9) *Either John loves Mary or Claire loves Robert.*

The information content of the two disjuncts can be represented by:

$$\ll loves, John, Mary; 1 \gg$$

and

$$\ll loves, Claire, Robert; 1 \gg,$$

respectively. We also have an operation to represent the disjunction.

$$\ll loves, John, Mary; 1 \gg \vee \ll loves, Claire, Robert; 1 \gg$$

but we have, as yet, no way to determine how the agent confronted with the three infons:

$$\ll loves, John, Mary; 0 \gg$$
$$\ll loves, Claire, Robert; 0 \gg$$
$$\ll loves, John, Mary; 1 \gg \vee \ll loves, Claire, Robert; 1 \gg$$

is to infer a contradiction. This may seem easy for the atomic infons, but not for complex infons, since negation is not defined. The difficulty can be overcome by letting situations branch according to the structure of Beth-tableaus, so each disjunctive piece of information would open two new branches, one for each disjunct.

Let σ be an information state and $\xi \vee \rho$ a disjunctive infon, and let

$$\sigma' = \sigma \oplus (\xi \vee \rho)$$

where

$$\sigma' = \sigma \oplus << \mathit{branch\text{-}of}, \sigma', \sigma_1; 1 >> \wedge \\ << \mathit{branch\text{-}of}, \sigma', \sigma_2; 1 >> \wedge \bigwedge \Xi)$$

and

$$\sigma_1 = \sigma \oplus \xi \text{ and } \sigma_2 = \sigma \oplus \rho$$

and where $\bigwedge \Xi$ is the conjunction of all infons supported by both σ_1 and σ_2 but not by σ.

3.3. Some meta-notions

To handle the most interesting type of composite infons, namely the ones with restrictions, we first have to define few notions, i.e., consistency, coherence, coherent extensions, licensed extensions, and minimal coherent extensions (MCE for short) of situations.

3.1. A situation σ is inconsistent iff there is an atomic infon ξ for which both $\sigma \models \xi$ and $\sigma \models \neg\xi$ hold. A situation is consistent if it is not inconsistent.

3.2. An information state s is called incoherent iff there is a restricted situation theoretic object, $\eta \mid p$ such that $\eta \mid p$ has an occurrence in s and $p \in EXT(FALSE)_\sigma$

3.3. An information state s is coherent iff for any restricted situation theoretic object, $\eta \mid p$ such that $\eta \mid p$ has an occurrence in s and $p \in EXT(TRUE)_\sigma$.

3.4. An information state s is incomplete iff it is neither incoherent nor coherent. Informally speaking a situation is incomplete if no restrictive proposition is false in it, but some are not true either.

3.5. A situation σ' is a coherent extension of the situation σ if for any infon ξ $\sigma \models \xi$ implies $\sigma' \models \xi$ and σ' is coherent.

3.6. A situation σ' is a licensed extension of the situation σ if for any infon ξ $\sigma \models \xi$ implies $\sigma' \models \xi$ and $\sigma \not\models \xi$ and $\sigma' \models \xi$ implies that

1) if it is a disjunction based extension, i.e., there exists a branch of σ say σ'' such that $\sigma'' \models \xi$ or

2) if it is a lexicon based extension, i.e., there exists an LCO-chain $LCO_0(\alpha_o), \ldots, LCO_n(\alpha_n)$ such that $\alpha_o \in LEX(\sigma)$ and $\alpha_{i+1} \in LEX(LCO_i(\alpha_i)$ and $p \in EXT(TRUE)_{LCO_n(\alpha_n)}$ and $LCO_n(\alpha_n) \models \xi$.

3.7. MCE is defined in terms of the partial ordering "smaller".

1) A disjunction based licensed extension of a situation is a smaller one than a lexicon based licensed extension.

2) A licensed extension σ' of the situation σ is a smaller extension than a further extension σ'' of σ' .

3) $MCE(\sigma)$ is the smallest coherent licensed extension of σ.

3.4. Restrictions

In this section we describe what to do with restrictions in the update procedure.

• If a restriction p of the infon ξ is true in the initial situation σ, the result of updating σ with the p-restricted infon $\xi \mid p$ is the same as an update with the unrestricted infon ξ. For a coherent situation σ $\sigma \oplus (\xi \mid p) = \sigma \oplus \xi$ provided that $\sigma \oplus (\xi \mid p)$ is coherent.

• If a restriction p of the infon ξ is false in the initial situation σ, the result of updating σ with the p-restricted infon $\xi \mid p$ is the absurd situation. If $\sigma \oplus (\xi \mid p)$ is incoherent then $\sigma \oplus (\xi \mid p) = \otimes$.

• If $\sigma \oplus (\xi \mid p)$ is incomplete, we **try** to accommodate p.

•• p can be accommodated to σ iff there exists a licensed coherent extension of $\sigma \oplus (\xi \mid p)$.

••$_1$. The result of the accommodation is: $(\sigma \oplus \xi) \mid p)$ if the accommodation is possible and the absurd information state $\otimes$ otherwise.

Note, that this definition of accommodation corresponds to global accommodation.[1] Whatever restrictions we accommodate this way, it restricts any

[1] For the distinction between global and local accommodation see Goldberg–Kálmán–Szabó 1991.

further continuation (i.e., extension, based on further updates) of the situation too. A **theory** that handles both global and local accommodation goes far beyond the limits of the present paper. For that reason we only hint to what local accommodation would look like.

$\bullet\bullet_2$. The result of the accommodation is: $(MCE(\sigma \oplus (\xi \mid p))$ if the accommodation is possible and the absurd information state $\otimes$ otherwise.

4. Testing deferred pieces of information

In this section we show how an update semantics can be given for a first order language with presuppositions. Furthermore to see the inferential behaviour of deferred pieces of information we extend the language with two test operators.

We start with the language of classical first order logic (LOFOL for short). To express presuppositions we build up a new language PRL. The rule that generates formulae of PRL is the following:

$$\phi_1 \in FORM_{LOFOL} \text{ and } \phi_2 \in FORM_{LOFOL} \text{ iff } (\phi_1, \phi_2) \in FORM_{PRL}$$

Both test operators **N** and **P** are one-place sentential operators. Their intuitive meanings are 'necessarily' and 'presumably' respectively. We assume that these test operators occur only as the outmost operator operating on a LOFOL formula.

To give an update semantics for PRL is straightforward. In case of a formula $(\phi_1, \phi_2) \in FORM_{PRL}$ and a situation σ,

$$\sigma[\![(\phi_1, \phi_2)]\!] = \sigma \oplus [\phi_1] \mid [\phi_2]$$

where $[\phi_1]$ is the infon associated with the formula ϕ_1, and $[\phi_2]$ is a proposition associated to the formula ϕ_2.

We associate infons to ordinary sentences and propositions to presuppositions, to reflect the fact the presuppositions are situated. They supposed to be true for example either in all coherent continuations of the updated situation, or in the minimal coherent extension of the updated situation, etc. The representation of ordinary sentences is turned into propositions by the update process. We get the proposition:

$$(\sigma \oplus \xi) : (\lambda x[x \models \xi])$$

The update conditions for the sentences of prefixed with the two test operators are the following:

$$\sigma[\![\mathbf{N}(\phi)]\!] = \begin{cases} \sigma & \text{if } |[\phi]| \in EXT(TRUE)_{\sigma'} \\ & \text{for any coherent extension of } \sigma \\ \otimes & \textit{otherwise} \end{cases}$$

and

$$\sigma[\![\mathbf{P}(\phi)]\!] = \begin{cases} \sigma & \text{if } |[\phi]| \in EXT(TRUE)_{MCE(\sigma)} \\ \otimes & \text{otherwise} \end{cases}$$

It is easy to check that with these update conditions for all globally accommodated presuppositions it follows that their necessitated version gives a positive result after the accommodation, i.e. keeps the situation unchanged. As for the locally accommodated presuppositions as well as for the implicit information pieces only the version prefixed with presumably goes through immediately after the accommodation. Later even that might vanish.

We argued elsewhere extensively that these are the desired results at least from a logical point of view (see for example Pólos–Masuch 1995). We now illustrate the argument with an example, which is a slight modification of one in Goldberg–Kálmán–Szabó (1991): *The jury investigates the photo of the 200 m race. The winner smiles on the picture.*

The second sentence of this discourse presupposes the existence of a unique winner. On the basis of these two sentences we assume that the race was not a dead heat. The winner is found in on one branch (opposite to the dead heat) of the LCO of the race.

What the next discourse shows is that we cannot cancel what actually presupposed was. In this case the existence of the unique winner, but we may well find that that race was a dead heat and the unique winner was the winner of another, more specifically mentioned race: *The jury investigates the photo of the 200 m race. The winner smiles on the picture. He got enough advantage in the first nine disciplines, and with this draw he won the pentathlon.*

And indeed the (global) accommodation procedure collects the restrictions as constraints on the coherent closure of the end result of the update procedure. The deferred pieces of information provided the licence and made (as a side effect of the update procedure) some implicit pieces of information available. In case of a continuation of that discourse the licence for the accommodation is still required, but it might well be a different, more powerful authority (more specific LCO situation) that provides the licence this time. And that changes the status of the implicit pieces of information too.

5. Conclusion

In this paper we sketched a situp semantics for a minimal language with presuppositions. Our semantics took advantage form the way in which situation theory handles restrictions. This semantics followed the general intuitions about the role of deferred pieces of information in presupposition accommodation. We limited the study of the inferential behaviour of deferred information pieces to the semantics of two relevant test operators.

Our ultimate conclusion was that this inferential behaviour of deferred pieces of information, at least through the looking glass of our semantics is in a large extent non-monotonic. László Kálmán believes that their behaviour is monotonic, and gave a framework to show how this is possible (Kálmán 1990). I believe that inferencing from deferred information pieces is (unfortunately) non-monotonic and sketched a semantics to show how that is possible. There should be something valuable about the notion of deferred information if it seems to be useful from both of these so utterly different perspectives.

References

Aczel, P. 1989. Non-well Founded Sets. CSLI Lecture notes.

Aczel, P.-Lunnon, R. 1991. Universes with Parameters. In: Barwise, J. *et al.* (eds): Situation Theory and Its Applications. Vol. 2. CLSI.

Asher, N. 1986. Belief sentences in DTR. JPhL.

Asher, N. 1993. Reference to Abstract Objects in Discourse. Kluwer, Dordrecht, Holland.

Barwise, J. 1983. Situation in logic, 1. CSLI report.

Barwise, J. 1985a. Situation in logic, 2. Conditionals and conditional information. CSLI report.

Barwise, J. 1985b. Situation in logic, 3. Situations sets and axiom of foundation. CSLI report.

Barwise, J. 1988. Situation in logic, 4. CSLI report.

Barwise, J.-Cooper, R. 1994. Simple situation theory and its graphical representation. DYANA deliverable R2.1C.

Barwise, J.-Etchemendy, J. 1986. The Liar. Oxford University Press.

Barwise, J.-Moss, L. 1991. Saarbrücken Lecture Notes.

Barwise, J.-Perry, J. 1983. Situations and Attitudes. MIT Press.

Devlin, K. 1989. Information and Inference. MIT Press.

Goldberg, J.–Kálmán, L.–Szabó, Z. 1991. Internal and external presuppositions. Presentation on the Diana Workshop on Presuppositions in Nijmegen.

Heim, I. 1982. The semantics of definite and indefinite noun phrases. Ph.D. dissertation. Mass.

Heim, I. 1983. The familiarity theory of definites. In: Bäuerle *et al.* (eds): Meaning, Use and Interpretation of Language. De Gruyter, Berlin.

Kamp, H. 1981. A Theory of Truth and Semantic Representation. In: Groenendijk, J. *et al.* (eds): Formal Methods in the Study of Language. Mathematical Centre, Amsterdam.

Kamp, H.–Reyle, U. 1993. From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory. Kluwer, Dordrecht, Holland.

Kálmán, L. 1990. Deferred information: semantics of commitment. In: Kálmán, L.–Pólos, L. (eds): Papers from the Second Symposium on Logic and Language. Akadémiai Kiadó, Budapest.

Kálmán, L.–Pólos, L.–Szabó, Z. 1989. Can representations solve problems? ALL Technical Report.

Kálmán L.–Szabó, Z. 1990. D.I.R.T. An overview in Stokhof, M.–Torenvliet, L. (eds): Proceedings of the Seventh Amsterdam Colloquium. Amsterdam.

Landman, F. 1986. On pegs and alecs. In: Landman, F.: Towards a theory of information. Ph.D. dissertation. University of Amsterdam.

Landman, F. 1987. A handful version of DRT. Ms.

Lunnon, R. 1991. Generalized universes. Ph.D. thesis. University of Manchester.

Partee, B. 1984. Nominal and temporal anaphora. L&Ph.

Pólos, L. 1987. Structured domains in DRT (Typed or type free?). In: Ruzsa, I.–Szabolcsi, A. (eds): Proceedings of Symposion on Logic and Language. Akadémiai Kiadó, Budapest.

Pólos, L.–Masuch, M. 1995. Information states in situation theory. In: Pólos, L.–Masuch, M. (eds): Applied Logic: How What and Why. Kluwer, Dordrecht, Holland.

Tarski, A. 1956. Logic, Semantics, Metamathematics. Oxford University Press.

van Eijck, J. 1985. Quantification in natural language. Ph.D. dissertation. Rijkuniversiteit, Groningen.

Veltman, F. 1989. Update Semantics. Ms.

Veltman, F. 1995. Defaults in update semantics. Forthcoming.

Address of the author: László Pólos
Department of Symbolic Logic
Eötvös Loránd University, Budapest
Hungary
and
CCSOM, University of Amsterdam
Sarphatistraat 143
1018 Amsterdam
The Netherlands
e-mail: laszlo@ccsom.uva.nl

Acta Linguistica Hungarica, Vol. 42 (3–4), pp. 325–339 (1994)

EXISTENCE PRESUPPOSITIONS AND CATEGORY MISTAKES

CHRIS FOX

1. Introduction

If we are interested in the truth conditional semantics of natural language, our theories should say something about syntactically well formed expressions which appear to be semantically infelicitous.

This paper is concerned with two such examples of infelicity, both of which involve false existence presuppositions, one with a non-denoting definite descriptor, the other with a non-denoting anaphor:

(1) *The present queen of France is bald.*

(2) *Every man walked in. He whistled.*

Both of these examples seem to presuppose the existence of terms which do not exist. Even the negation of these discourses require the existence of these terms. In general, the problem to be addressed here is the status of such examples in truth conditional semantics.

One possibility is to use a notion of truth-value gap (van Fraassen 1966), which would allow us to say that propositions representing sentences, and discourse, can be neither true nor false. This raises the question of whether the truth-value gaps should be something expressed in the representation, or something about the representation and its models. In the former case, we might adopt a non-classical three-valued logic. In the latter case, we assume that the logic is incomplete with respect to certain assertions. In both cases, we seem to be saying that (the representation of) infelicitous sentences are the kinds of objects to which truth values can be ascribed.

However, there is an alternative that appears in some theories which address the logical paradoxes, such as the liar: *this sentence is false*. This is a well formed sentence, yet any classical theory which gives it a truth value becomes inconsistent.

1216–8076/94/$ 5.00

Many formalisms brush this problem under the carpet, and use strong typing in order to prevent the self-application present in such examples. Exception to this are theories of truth, propositions and properties (often referred to, collectively, as Property Theory). Such theories take seriously the idea that propositions should be considered to be primitive objects, whose truth conditions can be expressed by means of a truth predicate.

In general, the presence of a truth predicate gives rise to logical paradoxes in formal theories. Property Theory can avoid these paradoxes by effectively limiting the application of the truth predicate to only non-problematic terms. In Turner's Property Theory (PT) which axiomatises Aczel's Frege Structures (Turner 1992; Aczel 1980), the axioms of truth governing the truth predicate can only be applied to those terms which we can prove to be propositions. The axioms governing propositionhood do not allow a proof that paradoxical expressions are propositions, despite their appearance.

The paradoxes are thus taken to be examples of a "category mistake"; although they have the form of a proposition, it is a mistake to assume that their semantic representation can be classified as such.

It seems to me that the same argument can be applied in the case of sentences with false existence presuppositions: although they have the appropriate form, we might argue that it is a category mistake to assume that their semantic representation must be taken to be a proposition. Felicity is thus equated with propositionhood.

2. Property Theory

2.1. The basic theory

The PT consists of a language of terms and a language of well formed formulae. The terms are those of the untyped $\lambda\beta$-calculus, together with logical constants. Some of the terms correspond with propositions. Propositions that are combined appropriately with the logical constants form new propositions. Properties are terms that form propositions when applied to another term. Some propositions will be true propositions. The truth of non-atomic propositions depends upon the truth of the constituent propositions and the nature of the logical connectives that combine them.

The notions of proposition and true proposition are not defined syntactically (this would lead to the logical paradoxes, as the terms are untyped), rather, they are characterised by axioms in meta-language: the language of well formed formulae (wff). This is a first order language with two predicates, P and T, that characterise those terms which are propositions and true propositions respectively.

2.1.1. The formal theory

The following presents a formalisation of the languages of terms and wff, together with the axioms that provide the closure conditions for P and T.

The language of terms

Basic vocabulary:

Individual variables:	$x, y, z, \ldots$
Individual constants:	$c, d, e, \ldots$
Logical constants:	$\vee, \wedge, \neg, \Rightarrow, \Xi, \Theta$

Inductive definition of terms:

(i) Every variable or constant is a term.

(ii) If t is a term and x is a variable then $\lambda x.t$ is a term.

(iii) If t and t' are terms then $t(t')$ is a term.

The language of wff

Inductive definition of wff:

(i) If t and s are terms then $s = t, \mathrm{P}(t), \mathrm{T}(t)$ are atomic wff.

(ii) If φ and φ' are wff then $\varphi \,\&\, \varphi'$, $\varphi \vee \varphi'$, $\varphi \to \varphi'$, $\sim\varphi$ are wff.

(iii) If φ is a wff and x a variable then $\exists x \varphi$ and $\forall x \varphi$ are wff.

The theory is governed by the following axioms:

Axioms of the $\lambda\beta$-calculus

$$\begin{aligned} \lambda x.t &= \lambda y.t[y/x] \; y \text{ not free in } t \\ (\lambda x.t)t' &= t[t'/x] \end{aligned}$$

This defines the equivalence of terms.

The closure conditions for propositionhood are given by the following axioms:

Axioms of propositions

(i) $\mathrm{P}(t)\,\&\,\mathrm{P}(s) \rightarrow \mathrm{P}(t \wedge s)$

(ii) $\mathrm{P}(t)\,\&\,\mathrm{P}(s) \rightarrow \mathrm{P}(t \vee s)$

(iii) $\mathrm{P}(t)\,\&\,(\mathrm{T}(t) \rightarrow \mathrm{P}(s)) \rightarrow \mathrm{P}(t \Rightarrow s)$

(iv) $\mathrm{P}(t) \rightarrow \mathrm{P}(\neg t)$

(v) $\forall x \mathrm{P}(t) \rightarrow \mathrm{P}(\Theta\lambda x.t)$

(vi) $\forall x \mathrm{P}(t) \rightarrow \mathrm{P}(\Xi\lambda x.t)$

(vii) $\mathrm{P}(s \approx t)$

Truth conditions can be given for those terms that are propositions:

Axioms of truth

(i) $\mathrm{P}(t)\,\&\,\mathrm{P}(s) \rightarrow (\mathrm{T}(t \wedge s) \leftrightarrow \mathrm{T}(t)\,\&\,\mathrm{T}(s))$

(ii) $\mathrm{P}(t)\,\&\,\mathrm{P}(s) \rightarrow (\mathrm{T}(t \vee s) \leftrightarrow \mathrm{T}(t) \mathrm{v} \mathrm{T}(s))$

(iii) $\mathrm{P}(t)\,\&\,(\mathrm{T}(t) \rightarrow \mathrm{P}(s)) \rightarrow (\mathrm{T}(t \Rightarrow s) \leftrightarrow \mathrm{T}(t) \rightarrow \mathrm{T}(s))$

(iv) $\mathrm{P}(t) \rightarrow (\mathrm{T}(\neg t) \leftrightarrow\ \sim\mathrm{T}(t))$

(v) $\forall x \mathrm{P}(t) \rightarrow (\mathrm{T}(\Theta\lambda x.t) \leftrightarrow \forall x \mathrm{T}(t))$

(vi) $\forall x \mathrm{P}(t) \rightarrow (\mathrm{T}(\Xi\lambda x.t) \leftrightarrow \exists x \mathrm{T}(t))$

(vii) $\mathrm{T}(t \approx s) \leftrightarrow t = s$

(viii) $\mathrm{T}(t) \rightarrow \mathrm{P}(t)$

The last axiom states that only propositions may have truth conditions.

Note that the quantified propositions $\Theta\lambda x.t$, $\Xi\lambda x.t$ can be written as $\Theta x(t)$, $\Xi x(t)$, where the λ-abstraction is implicit.

This basic theory is very weak. The general approach for analysing semantic phenomena is to amend the theory either definitionally, or by strengthening it with more axioms and primitive notions. This is, of course, in addition to obtaining appropriate representations for natural language phrases.

It is straightforward to define some simple types. The notions of n-place relations can be defined recursively:

(i) $Rel_0(t) \leftrightarrow \mathrm{P}(t)$

(ii) $Rel_n(\lambda x.t) \leftrightarrow Rel_{n-1}(t)$

As can be seen, types in this theory correspond with predicates. We can write $Rel_1(t)$ as $\mathrm{Pty}(t)$.

Types can be declared that correspond with the notions of quantifier and determiner in Montague's IL (Turner 1992). If Q, R are types, then the functional type $(Q \Longrightarrow R)$ can be defined with:

$$(Q \Longrightarrow R)(g) =_{\mathrm{def}} \forall x(Q(x) \rightarrow R(gx))$$

3. The liar

The treatment of existence presuppositions that is proposed here suggests that there are some sentences that should not be taken to constitute propositions. This is the approach taken by PT towards logical paradoxes, typified by expressions such as:

This sentence is false.

In a theory with untyped abstraction, such as PT, a formal analogue of this paradox is given by:

$$\mathrm{RR}$$

where $R = \lambda x.\neg xx$. So R is a term that expresses the notion that its argument does not hold of itself. The paradox arises if we are in a position to consider the truth conditions of RR; with one application of $\lambda\beta$-reduction we can show that RR is equal to $\neg RR$:

$$\begin{aligned} \mathrm{RR} &= (\lambda x.\neg xx)(\lambda x.\neg xx) \\ &=_\beta \neg(\lambda x.\neg xx)(\lambda x.\neg xx) \\ &= \neg RR \end{aligned}$$

If we consider RR to express a proposition, then the theory is paradoxical. Other semantic theories, such as Montagues IL, avoid such paradoxes by using strong typing to prevent self-application. In PT, the paradox is avoided because we cannot prove RR is a proposition; we can only prove it forms a proposition when given a property, but we cannot show that R is a property. To tie this in with the notation used in PT, if we could show that RR is true:

$$\mathrm{T}(RR)$$

then we would be able to show that it is also false, by the axioms of truth and the equivalence between RR and $\neg RR$:

$$\sim\mathrm{T}(RR)$$

However, we can only consider the truth of RR, and apply the axioms of truth, if we can show that RR is a proposition:

$$\mathrm{P}(RR)$$

but we cannot.

The important point to note is that, in PT, propositionhood is not determined merely by the syntactic appearance of an expression: terms may have the appearance of propositions, but if they do not satisfy the requirements expressed by the axioms then we cannot consider them to be such, and we are not in a position to question their truth.

4. Definite descriptors in Property Theory

As we have seen, in PT, not all expressions represent propositions. This is an essential aspect of its treatment of paradoxical terms. The notion of propositionhood may be used to model a more general notion of felicity in discourse. We can set up the axioms in such a way that sentences whose presuppositions are not met cannot be proven to be propositions. This can be illustrated with definite descriptors. We can define a class Δ of natural language denotable individuals, and a class Pty_Δ of natural language denotable properties. Given p in Pty_Δ and s in Δ, then we can prove that ps is a proposition by adding the following axiom:

$$(\mathrm{Pty}_\Delta p \,\&\, \Delta s) \rightarrow \mathrm{P}(ps)$$

We may require that a definite descriptor σxqx (the x such that $\mathrm{T}(qx)$) is only provably in Δ if q has an extension in Δ:

$$\exists y(\Delta y \,\&\, \mathrm{T}(qy)) \rightarrow \Delta(\sigma xqx)$$

Thus, when it comes to evaluating sentences such as *the present queen of France is bald*, we cannot prove that the representation of the sentence is a proposition. This is because we cannot prove that *the present queen of France* is a natural language denotable, as there is no present queen of France. The failure to prove the propositionhood of a sentence means that we cannot apply the axioms of T to determine the truth conditions of the sentence; the sentence is not the sort of object whose truth conditions should be considered.

This can be generalised to plurals and mass terms by replacing this axiom with:

$$\forall x(\mathrm{T}(qx) \rightarrow \Delta x) \,\&\, \exists y(\Delta y \,\&\, \mathrm{T}(qy)) \rightarrow \Delta(\sigma x q x)$$

which says that a definite descriptor is denotable if the associated property is a property of denotables, and there is a denotable in its extension. The form of this axiom is justified in Fox (1993).

Notice that with definite descriptors, what is referred to as "accommodation" could be modelled by some form of abduction in the representation. Also, the axioms for P and T with material implication cope with the so-called "projection problem" for conditionals: existence presuppositions arising from the consequent are not projected to the top level if they are satisfied by the truth of the antecedent. To cope with the projection problem with disjunction requires some alterations to the axioms (Fox 1993).

5. Anaphora in Type Theory

Propositionhood can also be used to capture the notion of felicitous discourse in examples where there are singular pronouns lacking an antecedent. Indeed, this is embodied in the constructive-type theoretic approach to natural language semantics (Ranta 1991; Davila-Perez 1994).

In brief, such approaches take sentences in a discourse to provide the types, or specifications, of the objects which satisfy them. We might see parallels with Discourse Representation Theory (DRT) (Kamp 1981), where natural language provides type specifications (the conditions) for discourse objects. In DRT, we might say that a discourse is true if we can anchor the discourse markers in a way which satisfies the conditions. The constructive-type approach is different in that objects (proofs or witnesses) must also be produced for the conditions themselves.

In effect, objects (witnesses) which satisfy a discourse are passed on to the next sentence as an argument. These witnesses are either objects structured as pairs (which arise from indefinites, relative clauses, and sequences of sentences) or function types, which are specified by the representation of universal quantification and conditional constructions.

Singular pronouns are resolved by representing them with selector functions, which can pick elements from nested pairs.

Universal quantification is usually taken to block singular anaphoric reference, as in the example:

(3) Every man walked in. He whistled.

In DRT this is achieved by "box-splitting". In Dynamic Predicate Logic (Groenendijk–Stokhof 1991) it is achieved by stipulating that universal quantification is "externally static". In the constructive type interpretation, it is blocked because the first sentence gives rise to a function, not a nested pair. Hence, selector functions cannot be applied to this without producing an ill-formed expression.

One theory of types that has used for anaphora in the semantics of natural language is Martin-Löf's Type Theory (MLTT) (Martin-Löf 1982, 1984; Sundholm 1989; Ranta 1991; Davila-Perez 1994).

In MLTT, there are types and elements (or witnesses) of types. If T is a type, then:

$$w \in T$$

says that w is an element of that type.

Types can be said to correspond with propositions, and elements of a type with proofs of that proposition. This is a form of intuitionistic logic; a proposition is true if we can produce a proof of it. If the class of specified proofs is non-empty (or inhabited), then the proposition is true.

Various sorts of complex types can be defined in MLTT. Those of interest for donkey sentences and intersentential anaphora are the dependent types. These allow us to treat the witness for a type as a context for subsequent types. Taking the type operator Σ as a relevant example:

$$\Sigma x \in f.g$$

is a type formed from the types f, g. Viewed as a proposition, it is true if we can find an h such that:

$$h \in (\Sigma x \in f.g)$$

Witnesses of this type are pairs, so $h = \langle a, b\rangle$. The definition of types of this form in MLTT requires that if:

$$\langle a, b\rangle \in (\Sigma x \in f.g)$$

then it must be the case that:

$$\begin{array}{l} a \in f \\ b \in g[a/x] \end{array}$$

That is, h consists of a proof of f and a proof of g, where the proof of f has been substituted for all free occurrences of x in g. Here we see the context dependence created by the Σ type: the proof required for g depends upon the proof give for f. If the proof of f in effect contains discourse referents, then these become available in g.

In natural language semantics, the Σ operator can be used to represent indefinites, sentential conjunction, and relative clauses. Using Σ for indefinites means that witnesses to nominals (discourse referents in DRT) are made available to later parts of the sentence. Using Σ for sentential conjunction means that these discourse referents are also made available to subsequent sentences.

Sequences of the selector functions $\mathsf{fst}, \mathsf{snd}$ can be used to recourse through the nested pairs that result from discourse. The act of resolving the reference of a singular pronoun is achieved by replacing the occurrence of the pronoun in the semantics by selector functions operating on an appropriate argument.

Skipping the details of the translation procedure, the discourse:

(4) A man walked in. He whistled.

would be represented by:

$$\Sigma y \in (\Sigma x \in \text{man}'\text{walked-in}'x)(\text{whistled}'(\text{he}_0))$$

This MLTT proposition will be true if there is a witness of the form:

$$\langle\langle m, \varphi\rangle, \psi\rangle$$

where m is a man, and φ is a proof that that man walked in, and ψ is a proof that *He* whistled. More specifically, following the rules for Σ, ψ is a proof of $(\text{whistled}'(\text{he}_0))$ where all occurrences of y (if any) are replaced by

$\langle m, \varphi \rangle$. We can make the type specify that it was the previously mentioned man that whistled by replacing the representation of the pronoun with $\mathsf{fst}(y)$. The type expression for the consequent sentences then becomes equivalent to whistled$'(m)$. That is, the discourse is true if:

$$\langle \langle m, \varphi \rangle, \psi \rangle \in \Sigma y \in (\Sigma x \in \text{man}'\text{walked-in}'x)(\text{whistled}'(\text{he}_0))$$

where

$$\begin{aligned} m &\in \text{man}' \\ \varphi &\in \text{walked-in}'(m) \\ \psi &\in \text{whistled}'(\text{he}_0)[\langle m, \varphi \rangle / y] \end{aligned}$$

As already mentioned, the singular pronoun he_0 is resolved by replacing it with $\mathsf{fst}(y)$, so that:

$$\begin{aligned} \psi &\in \text{whistled}'(\mathsf{fst}\langle m, \varphi \rangle) \\ &\in \text{whistled}'(m) \end{aligned}$$

For universal quantification and conditionals, we need to make use of another dependent type operator, Π. The expression:

$$\Pi x \in f.g$$

holds of functions which take all proofs, or witnesses, w of f to a proof of $g(w)$. This is similar to universal quantification (or implication). It can be used to represent the universal determiner *all*, and *if ... then ...* constructions. As with box-splitting in DRT, Π blocks singular anaphoric reference. The sentence:

(5) If a man owns a donkey, he beats it.

can be represented by:

$$\Pi x \in (\Sigma y \in \text{man}'.(\Sigma z \in \text{donkey}'\text{own}'zy).(\text{beat}'(\text{it}_0)(\text{he}_0)))$$

The sentence is true if we can find a function that maps objects of the form:

$$\langle f, \langle d, \varphi \rangle \rangle$$

where f is a farmer, d is a donkey, and φ is a proof that f owns d, into a proof of:

$$\text{beats}'(\text{it}_0, \text{he}_0)$$

The anaphora are resolved if we replace them with selectors as follows:

$$\text{beat}'(\mathsf{fst}(\mathsf{snd}(x)))(\mathsf{fst}(x))$$

where x becomes instantiated with $\langle f, \langle d, \varphi \rangle \rangle$:

$$\text{beat}'(\mathsf{fst}(\mathsf{snd}(\langle f, \langle d, o \rangle \rangle)))(\mathsf{fst}(\langle f, \langle d, o \rangle \rangle))$$

which is equivalent to:

$$\text{beat}' df$$

For a compositional analysis, this is not the whole story. For more details, the interested reader is referred to the expositions of natural language semantics in MLTT by Ranta (1991) and Davila-Perez (1994).

In MLTT, our infelicitous example:

(6) Every man walked in. He whistled.

is represented by something like:

$$\Sigma y \in (\Pi x \in \text{man}' \text{walked-in}' x)(\text{whistled}'(\text{he}_0))$$

The pronoun in the representation of the second sentence cannot be resolved by replacing it with a selector as the previous sentence supplies a function, not a nested pair. The formation rules in MLTT do not support this representation as it would embody a type mismatch. As can be seen, this approach to discourse anaphora has parallels with the treatment of non-denoting definite descriptors given above.

Perhaps some objections to the use of MLTT for natural language semantics should be noted. Firstly, it is a non-classical theory. Witnesses must be provided for both properties and propositions. In this sense, the theory conflates the notions of property and proposition.

Additionally, there is no means of distinguishing false propositions; and there is no extensional-intensional contrast in the basic theory. Both are required in natural language semantics. (These drawbacks might be overcome by adding various "Universes" to MLTT, but this seems to admit the primitive nature of propositions, and so defeats the anti-realist arguments that favour MLTT over some more classical theory.)

To remedy these last two defects, and to ellucidate the connection with the treatment of non-denoting definite descriptors that I have proposed, we can implement MLTT in PT, as done explicitly by Turner (1992) and Smith (1984). Indeed, working with MLTT defined in PT gives some additional flexibility, and provides a classical notion of propositions and the standard connectives.

In principle, this provides the tools that are needed to give an analysis of anaphora with dependent types whilst keeping the classical distinction between properties and propositions (Fox 1994).

6. Type Theory in Property Theory

Types can be defined as properties in PT. We can give definitions for various operations on these types, such as intersection $\cap$, union $\cup$, difference $-$, cartesian product $\otimes$, disjoint union $\oplus$, and function space $\mapsto$ operators (Turner 1992). Only intersection and disjoint union will be illustrated here:

$$\begin{aligned} \cap &=_{\mathsf{def}} \lambda f.\lambda g.\{x : fx \wedge gx\} \\ \oplus &=_{\mathsf{def}} \lambda f.\lambda g.\{z : \\ &\qquad (\mathsf{fst}(z) \approx 0 \wedge f(\mathsf{snd}(z))) \\ &\qquad \vee (\mathsf{fst}(z) \approx 1 \wedge g(\mathsf{snd}(z))) \end{aligned}$$

where $\{x : t\}$ is syntactic sugar for $\lambda x.t$, especially when t is a property.

The definitions of type operators trivially support the theorems:

$$\begin{aligned} z \in (t \cap s) &\leftrightarrow z \in t \,\&\, z \in s \\ z \in (t \oplus s) &\leftrightarrow (\mathsf{fst}(z) = 0 \,\&\, \mathsf{snd}(z) \in t) \\ &\qquad \vee (\mathsf{fst}(z) = 1 \,\&\, \mathsf{snd}(z) \in s) \end{aligned}$$

where $x \in t$ is sugar for $\mathrm{T}(tx)$.

Pairs $\langle , \rangle$ and the selector functions fst, snd have their usual definitions:

$$\begin{aligned} \langle x, y \rangle &=_{\mathsf{def}} \lambda z.z(x)(y) \\ \mathsf{fst} &=_{\mathsf{def}} \lambda p.p\lambda xy.x \\ \mathsf{snd} &=_{\mathsf{def}} \lambda p.p\lambda xy.y \end{aligned}$$

so that:

$$\begin{aligned} \mathsf{fst}(\langle x, y \rangle) &=_\beta x \\ \mathsf{snd}(\langle x, y \rangle) &=_\beta y \end{aligned}$$

The dependent type operators Σ, Π of MLTT can be defined by:

$$\begin{aligned} \Pi &=_{\mathsf{def}} \lambda f.\lambda g.\{h : \Theta x(fx \Rightarrow gx(hx))\} \\ \Sigma &=_{\mathsf{def}} \lambda f.\lambda g.\{h : f(\mathsf{fst}(h)) \wedge g(\mathsf{fst}(h))(\mathsf{snd}(h))\} \end{aligned}$$

These definitions support the following theorems:

If $\mathrm{Pty}(f)$ and $\forall x(x \in f \rightarrow \mathrm{Pty}(gx))$ then:

$$\begin{gathered} \mathrm{Pty}(\Pi fg) \\ \mathrm{Pty}(\Sigma fg) \end{gathered}$$

and:

$$\begin{aligned} h \in \Pi fg &\leftrightarrow \forall x(x \in f \rightarrow hx \in gx) \\ h \in \Sigma fg &\leftrightarrow \mathsf{fst}(h) \in f \,\&\, \mathsf{snd}(h) \in g(\mathsf{fst}(h)) \end{aligned}$$

To paraphrase these definitions, $h \in \Pi fg$ requires that h is a function which takes an element (in MLTT terminology, a proof, or witness) of f and gives an element (proof) of g applied to that element of f. The expression $h \in \Sigma fg$ requires that h is a pair, where the first component of the pair is an element of f, and the second is an element of g applied to that element of f.

With both of these types, the evaluation of g is dependent upon the chosen element (proof) of f. In some sense then, the meaning of g depends upon the context created by f. This gives us the means to give an MLTT based treatment of anaphora and discourse in PT, as illustrated before.

Now we can adapt the approach given for the infelicitous definite descriptor to the infelicitous anaphoric example. Essentially, we can assign types to the basic terms that will prevent the infelicitous anaphoric example being shown to be a MLTT proposition (a PT property) when the pronoun is replaced by a selector function. Thus it is will not be possible to prove that the representation of the discourse forms a proposition when given a putative witness.

The appropriate types are as follows: denotable individuals will be in Δ, and MLTT properties will form PT properties when given a denotable, that is, they will be of the type $(\Delta \Longrightarrow \mathrm{Pty})$.

In the representation of the infelicitous example:

(7) Every man walked in. He whistled.

it is not possible to replace the pronoun with a selector that is provably a denotable (in Δ), so we cannot show that the discourse is of an appropriate type to have its truth conditions considered.

As mentioned before, the proposed treatment of the first, definite descriptor example uses the classical notion of proposition, while the second, anaphoric example uses constructive propositions (sentences are represented as types). Although there is insufficient space to give the details here, one way in which we can give a more uniform account is to provide a translation which turns constructive propositions into related classical propositions (Fox 1994).

7. Conclusions

Using Property-theoretic semantics, it is possible to use propositionhood to characterise felicitous discourse. To consider the truth conditions of discourse with: paradoxical expressions; non-denoting definite descriptors; and pronouns without an antecedent, is to make a category mistake.

It might be argued that if both the paradoxes and expressions with non-denoting nominals give rise to the absence of propositionhood, then we have conflated two very different phenomena. However, the failure to prove propositionhood arises for different reasons: with the paradoxes, it stems from failure to prove that a term is a property; with existence presuppositions it comes from failure to prove that a term is denotable. In the later case, an appropriate denotable term can be added, or accommodated. Propositionhood will then follow. In the former case, no consistent assumption can be added which will allow a proof of propositionhood.

This treatment of existence presuppositions is achieved in a first-order, two-valued classical theory, with fine-grained intensionality. As such it lends support to the view that this is all that is needed for the truth conditional semantics of a reasonably large fragment of natural language.

References

Aczel, P. 1980. Frege structures and the notions of proposition, truth and set. In: Barwise, J.–Keisler, X.–Keenan, E. (eds): The Kleene Symposium. North Holland Studies in Logic, 31–9. North Holland.

Davila-Perez, R. 1994. Constructive type theory and natural language. Computer Science Memorandum 206. University of Essex.

Fox, C. 1993. Plurals and mass terms in Property Theory. Ph.D. thesis. University of Essex, Colchester, U.K.

Fox, C. 1994. Discourse representation, type theory and property theory. In: Bunt, H.–Muskens, R.–Rentier, G. (eds): Proceedings of the International Workshop on Computational Semantics, 71–80. ITK, Tilburg, Holland.

Groenendijk, J.–Stokhof, M. 1991. Dynamic predicate logic. In: Linguistics and Philosophy 14: 39–100.

Kamp, H. 1981. Theory of truth and semantic representation. In: Groenendijk, J.–Janssen, T.–Stokhof, M. (eds): Formal Methods in the Study of Language. Mathematical Centre Tracts 135: 277–322. Amsterdam.

Martin-Löf, P. 1982. Constructive mathematics and computer programming. In: Cohen–Los–Pfeiffer–Podewski (eds): Logic, Methodology and Philosophy of Science VI, 153–79. North Holland.

Martin-Löf, P. 1984. Studies in Proof Theory (lecture notes). Bibliopolis, Napoli.

Ranta, A. 1991. Intuitionistic categorial grammar. In: Linguistics and Philosophy 14: 203–39.

Smith, J.M. 1984. An interpretation of Martin-Löf's Type Theory in a type-free theory of propositions. In: Journal of Symbolic Logic 49.

Sundholm, G. 1989. Constructive generalised quantifiers. In: Synthese 79: 1–12.

Turner, R. 1992. Properties, propositions and semantic theory. In: Rosner, M.–Johnson, R. (eds): Computational Linguistics and Formal Semantics. Studies in Natural Language Processing, 159–80. Cambridge University Press, Cambridge.

van Fraassen, B.C. 1966. Singular terms and truth value gaps. In: The Journal of Philosophy 63: 481–95.

Address of the author: Chris Fox
Department of Computer Science
University of Essex
Wivenhoe Park
Colchester
CO4 3SQ
United Kingdom
e-mail: foxcj@essex.ac.uk

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 with the following specifications:

operation system: DOS 2.0 or later;
floppy disk size: 5.25 or 3.5 inch;
floppy disk format: DS,DD (360/720 Kbyte) or DS,HD (1.2/1.44 Mbyte);
text editor: XYWrite, Word or Word Perfect;
text file: ASCII file without the formatting commands of the text editor;
special characters: choose some character combinations and use them consequently, for example: ə=\schwa{}; ɔ=\openo{}.

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, MTA Nyelvtudományi Intézet, H-1250 Budapest, P.O. Box 19.

Acta Linguistica Hungarica

AN INTERNATIONAL JOURNAL OF LINGUISTICS

Volume 42

Akadémiai Kiadó, Budapest
1994

ACTA LINGUISTICA HUNGARICA
VOLUME 42

CONTENTS

Book reviews

PRINTED IN HUNGARY
AKADÉMIAI NYOMDA, MARTONVÁSÁR