Editorial

The editors are pleased to welcome you to first issue of the sixth volume of FULL, an open access international journal providing a platform for linguistic research on modern and older Finno-Ugric or other Uralic languages and dialects. FULL publishes comparative research as well as research on single languages, including comparison of just Uralic languages or comparison across family lines. We welcome formal linguistic accounts as well as empirically oriented contributions.

The first article of this issue, by Andrea Deme, Katalin Gugán, Bálint Sass, and Katalin Mády is about the identification and measurement of implicit attitudes towards linguistic innovations, a key factor in language change. The authors argue that a method borrowed from social psychology, the Implicit Association Test, can be fruitfully adapted for the purpose. In the paper they put forward their version of the IAT, and test it on data from Hungarian in two experiments.

The second article, by Bernadett Bíró and Katalin Sipőcz, is a study of two ditransitive constructions in the Ob-Ugric language Mansi. They show that the two constructions are a special case of the cross-linguistically common alternation between indirective and secundative ditransitive syntax. The study demonstrates that the alternation is mainly determined by topicality.

The third article, by Merilin Miljan, Elsi Kaiser, and Virve-Anneli Vihman, concerns how semantic roles are assigned to arguments in sentence processing, with particular focus on the role of case. The language investigated is Estonian. They report an experiment exploring the complex interplay of case, animacy, and number in determining the semantic roles of arguments in the initial part of utterances that is processed before parsing the verb.

We take this opportunity to thank the anonymous reviewers who generously lent their time and expertise to FULL.

Our publications can be freely accessed and downloaded without any need for prior registration. At the same time, those who register, or have already registered, are provided with the benefit of getting notified of new issues, calls, etc. via email.

FULL welcomes manuscripts from all the main branches of linguistics, including phonology, morphology, syntax, semantics and pragmatics, employing a diachronic or synchronic perspective, as well as from first language acquisition and psycholinguistics. Whatever the theoretical or empirical orientation of the contributions may be, our leading principle is to maintain the highest international standards.

The Editors

1   Introduction

The following text is intended as supplementary material for the article ***Towards capturing implicit innovative language attitude using an auditory Implicit Association Test.*** It contains a more detailed structural description of the variables tested in Experiment 1,[1] including the results of frequency studies on the basis of corpus queries, which served as a background for the interpretation of the test results. Besides, the labels "more frequent" and "less frequent", which were assigned to the variants to be able to distinguish them systematically, are also based on the results of the corpus queries. However, it is also necessary to add that we did not mean to overestimate the significance of the search results, that is, we do not claim that this opposition between the forms faithfully represents "real" language use. Still, these simple labels will hopefully be useful for practical purposes (i.e., consistent differentiation of the two variants of a variable).

The test used for eliciting grammaticality judgments (and the possible evaluation differences between the different age groups) contained 11 structures that exhibit variation in Modern Standard Hungarian (abbreviated as MSH in the following). As many details of the history of Hungarian are thoroughly researched, most of these variants could be classified reliably as "innovative" or "conservative" on a diachronic basis. However, some of these variants have co-existed for hundreds of years, which would make it rather questionable to label them mechanically "innovations" or "archaisms" on the basis of their diachronic background within the framework of the present study investigating synchronic language use. In what follows, we will discuss each of the tested linguistic variables in more detail, with respect to the following guidelines: the structural description of the given form, its diachronic background, and data concerning the use of the variants, obtained on the basis of queries of the Hungarian National Corpus (HNC; for a brief description, see below). Hopefully, these different types of background information will help the interpretation of test data.

The new version of the HNC (Oravecz et al. 2014) is a large, modern corpus of contemporary Hungarian, equipped with thorough linguistic annotation and a fast and detailed corpus query interface. The annotation contains morphological information (inflection, derivation, compounding, morphemes) and phonological information (natural classes, CV skeleton) as well. The currently available part of it contains six genres in the following proportions: journalism 35.5%, literature 6.6%, (popular) science 20.1%, personal (web forums) 7.9%, official 16.9%, transcribed spoken (radio) 13.9%.[2] With respect to *modality* (speech vs. written text), the *speaking style* (or in other words, spontaneity, that is, the timing of speech planning processes, and the articulation of the utterances) and *register* (specifically the formal-informal continuum), the following features are considered to be characteristic of these subcorpora.

"Journalism" contains only written material that can be considered to be closer to the "formal" end of the continuum. "Literature" is also compiled only of written material and contains texts that, naturally, intertwine formal and informal registers. Similarly to "literature", "(popular) science" contains written texts, but more to the formal end of the

---

[1]   For the description of Experiment 1, see Section 3 in the article.

[2]   To the best of our knowledge, this subcorpus of HNC is the largest morphologically annotated corpus of spoken Hungarian.

formal-informal continuum. "Personal" is compiled of written texts (posts of web forums) that may be positioned, again, more to the informal end of the continuum. "Official" refers to transcribed speeches that were written in advance and read aloud at the Parliament (this is the bigger part of the subcorpus), but this also contains some spontaneous speech that was improvised on the spot. "Transcribed spoken (radio)" is a subcorpus in which mainly transcribed spontaneous speech and a smaller amount of read aloud speech can be found that was originally recorded on a country-wide radio station (Kossuth Rádió). The registers used in these texts may be considered mixed, again, but more to the end of the formal continuum: although informal utterances are also observable in this subcorpus, it is clearly noticeable that both the (editorial) staff and the guests are aiming at producing more elevated types of speech. It is important to note that simultaneously informal and spontaneous speech is essentially missing from the HNC, and this might be the cause of the underrepresentation of the supposedly innovative variants analyzed and investigated in the study. In what follows, we summarize the results of the corpus analysis, and we refer to the tested phenomena (listed in Section 3.1 of the article) in the heading of each subsection.

2   Article drop (phenomena 1, 2, 3)

The Hungarian definite article *a/az* emerged during the Old Hungarian period through the functional split of the distal demonstrative pronoun *az* 'that', a grammaticalization change that was attested in various languages (Heine & Kuteva 2004, Harris & Campbell 1995, Givón 2001). Its use displays gradual extension during Old Hungarian, appearing in more and more grammatical contexts where it marks definiteness (I. Gallasy 1991, 1992, Egedi 2014).

However, in the test we looked at examples which could (theoretically) be interpreted as a change in the opposite direction:[3] the article does not appear in environments in which it would be compulsory, or at least preferable in MSH. An example of this type is shown in (1), observed by Nádasdy (2008).[4]

(1) *Nyugatiba   bejössz* […]
      Western-ILL   drop.in.PRS.2SG
      'You come into the Western (Railway Station)'

With respect example (1), Nádasdy claims that while this form does not conform to MSH (and one reason for this is the lack of the article before *Nyugati)*, it is still generally observable in spoken language, and the environment that might trigger article drop is the sentence-initial position of the noun. In addition to the position of the noun, we decided to test for an additional variable, namely whether names of institutions like "Nyugati" in (1) show a tendency to be used without a definite article.

---

[3]   However, it is almost impossible to investigate whether in these specific instances the use of the article had already been general, and the drop of the article is a further change, or (more likely) there has always been a significant amount of variation, and, due to so far unknown reasons, the more archaic, article-less pattern started to spread in MSH.

[4]   ACC = accusative, DAT = dative, DEL = delative, ILL = illative, INF= infinitive, INS = instrumental, MOD = marker of epistemic modality; PL = plural, PRS = present, PST = past, PV = preverb (marking telicity), SUP = superessive.

Concerning sentence-initial position, the query of HNC ran with the following settings. First we searched for plural nouns in the nominative in sentence initial position, with a further restriction that the noun cannot be suffixed with possessive markers (example (2) below is a typical example of this pattern).[5]

(2) *(az/0) ajtók   a   bal   oldalon   nyílnak.*
    (the/0) door.PL   the   left   side.SUP open.PRS.3PL
    'Doors open on the left.'

The number of hits when these nouns occurred with and without the definite article was 7,480 and 14,430, respectively. For a more detailed analysis, we have chosen a random list of one thousand hits of the latter group. It turned out that the overwhelming majority of the article-less items were titles (appearing in newspapers, scientific works etc.), which do show a tendency to drop articles owing to the specific style they represent. The rest of the examples cannot get a uniform explanation: even if we could assume (on the basis of our Hungarian competence) that certain structures can have competing variants that are identical except for the presence of a determiner, the competition is not always between a definite determiner and a null, but can just as well be a competition between an indefinite determiner and a null. Therefore, as there are very few genuine examples of article drop,[6] we could not attest the general nature of sentence-initial article deletion on the basis of HNC, which, naturally, can be due to several reasons, including the fact that spontaneous informal language use is underrepresented in the corpus, as noted above.

As for names of institutions, in MSH these occur with the definite article (as an indirect evidence, the comprehensive handbook of normative linguistics explicitly states this, cf. Grétsy & Kovalovszky 1980, 104). However, on the basis of our own observations we had the impression that some speakers prefer to drop the article before these. To obtain a larger number of data (and, as in this case the hits did not need much manual analysis), we searched the entire corpus for all the instances for the verbs *bemegy* 'enter' and *beugrik* 'stop by' with an illative-marked noun starting with a capital letter in its immediate environment. The query resulted in 220 hits in which the noun had an article and no hits for article-less nouns. Therefore, the conclusion is similar to that of the previous case: article drop before names of institutions is not yet attestable in HNC, either because of the novelty of this phenomenon, or due to nature of the corpus.

A final test point was to see article use with nouns in sentence medial position. Naturally, this is too broad a category to be tested in general, therefore, we selected sentences to be tested in which we happened to observe article drop ourselves (e.g., the article-less version of example (3) was attested by us in spontaneous communication). In the test these sentences occurred with and without the article (see 3.1 in the article), and the focus of the corpus queries was finding the distribution of these two patterns.

---

[5] This restriction was necessary as possessive-marked nouns, which are predominantly interpreted as definite in themselves, differ significantly in this respect. In order to illustrate this, we looked for possessive marked nouns in sentence initial position as well. In this case we did not perform a manual analysis of the findings, but it is telling that there were 6619 hits with an article, and about 87,000 hits without an article.

[6] E.g., *"Kezdők a Standard Scant hajtsák végre, a haladóknak viszont lehetőségük van kizáró listát készíteni."* 'Beginners should execute standard scan, however, the advanced ones have the option to make a disqualifying list.'

(3) *Sajnos   (a)   kórussal   lesz   fontos   prób-ám.*
   Unfortunately (the) choir.INS   will.be.3SG important   rehearsal.1SG
   'Unfortunately, I will have an important rehearsal with (the) choir.'

The test itself aimed at investigating the use (and potential drop) of the definite article before nouns with and without possessive suffixes, therefore, we also tried to find parallels for the investigated structures from HNC. First, we looked for sentences similar to (3) in the corpus: those starting with the sentence adverbial *sajnos* 'unfortunately', followed by nouns with an instrumental marker, the latter either preceded by an article, or immediately following the sentence adverbial. There were 97 hits with the article and 67 without it (i.e., adjacency of the adverbial and the noun), but the manual analysis of these latter cases showed that none of these could be taken as examples of article drop (i.e., the noun in the instrumental could not have a determiner in the given context, or, if it could have some determiner, it would not necessarily be the definite article). Therefore, our observation cannot be confirmed by corpus data. Nevertheless, these findings justify referring to the artcle-less variant as "rare", similarly to the other hypothesized cases of article deletion.

In contrast, we looked at instances of nouns meaning 'sibling' *(öcs* 'younger brother', *báty* 'older brother', *húg* 'younger sister', *nővér* 'older sister') marked with a first person singular possessive suffix *(-m)* in HNC. The specific patterns we looked for was verb + definite article + possessive marked noun vs. verb + possessive-marked noun appearing adjacently (again, we had similar structures in the test). In this case, there were 481 hits for the pattern with the article, and 219 for the pattern without the article. The manual analysis of the latter group showed that there were 112 structures that can be considered as cases of article drop, i.e., when the noun could have a definite article (and, more than likely, it could only have a definite article), but the speaker chose not to use it. It seems to be the case that possessive-marked kinship terms can freely occur without the definite article, still, this pattern occurs less often than the other pattern.

3   Presence or absence of the subordinator *hogy* (phenomena 4, 5, 6, 7, 8)

The most general subordinator in Hungarian is *hogy* 'that', a complementizer that developed presumably during the Proto-Hungarian period through the reanalysis of the question word *hogy* 'how' in embedded questions or, alternatively, of the pronominal adverbial *hogy* 'as; the way that', which were homophonous at the time of this change (in MSH, the latter is *ahogy;* for a general description of the grammaticalization process, see Haader 1991). The investigations of the earliest sources reveal that by Old Hungarian the complementizer appears in all of those functions that it has in MSH (Haader 1995, Bácskai-Atkári & Dékány 2014). However, it was not an obligatory marker of finite subordination then, and its use is not obligatory in MSH, either. Summarizing the typical instances when the use of *hogy* is optional, Kenesei (1992, 673-679) distinguishes external and internal conditions of *hogy*-deletion. The former category covers conditions pertaining to the grammatical relationship of the main sentence and the subordinate sentence, while the latter category encompasses features that characterize the given subordinate sentences headed by the complementizer *hogy*.

Kenesei (1992, 674) points out that the external criteria of complementizer deletion are reducible to a single principle: deletion is only possible if the subordinate clause is properly governed by the verb of the matrix clause, i.e., the verb of the matrix clause

assigns theta-role to the subordinate clause, and the verb and the subordinate clause have to occupy sister nodes. For instance, this excludes cases when the matrix clause contains structural focus or negation, as in these cases the verb obligatory moves to a higher structural position, therefore it cannot be adjacent to the complementizer. The complementizer cannot be omitted in those cases when the subordinate clause precedes the main clause, either, and the reason behind this restriction is the same: the verb of the main clause and the complementizer of the subordinate clause cannot be adjacent in that pattern.

Concerning inner criteria, one of them (that overrides all principles listed above) is that the complementizer *hogy* can be omitted freely if the subordinate clause is an embedded question. Another criterion is the modality of the subordinate clause: if it is an embedded imperative, the complementizer can be deleted as well. Furthermore, if there is a second subordinate clause embedded into the subordinate clause headed by *hogy*, and, as a result of this, the two complementizers are adjacent, some complementizers facilitate the deletion of *hogy*. Finally, it is necessary to mention that there must be individual differences between the speakers concerning *hogy*-deletion, and stylistic rules can also influence this phenomenon. Most Hungarians learn at school that it is "not elegant" to iterate the complementizer *hogy* when embedded clauses follow each other, and in certain cases the written samples show omission even in cases when its grammaticality is dubious.

Out of the three sections that cover the presence or absence of *hogy* in the test (phenomena 4, 5-7, 8 in Section 3.1), one set contains examples where the complementizer could be freely omitted in MSH, as in example (4) below.

(4) *Azt*   *hittük,*   (*hogy*)   *itt*   *tilos*   *a*   *dohányzás.*
    that.ACC   think.PST.3SG<1PL   (that)   here   forbidden   the   smoking.
    'We thought (that) smoking was forbidden here.'

In order to acquire data that could facilitate the interpretation of the test results (presented in Section 3.3 of the article), there were three verbs (potentially subcategorizing for *hogy*-clauses) that were taken as the starting points of corpus query in HNC. We searched for the occurrences of the verbs *mond* 'say',[7] *hisz* 'think, believe', *bevall* 'admit' in the objective conjugation, as in MSH (as opposed to earlier periods) those verbs that properly govern the subordinate clause cannot be in the subjective conjugation. The search results were filtered for subcorpora, and a random sample of 250 items of the six subcorpora was generated for each of these three verbs, so altogether we looked at 1,500 occurrences of each verb. These were then analyzed manually for the feature investigated (overt complementizer, deleted complementizer, and irrelevant cases). Figure 1 shows the results of this analysis.[8]

The data reveal that there does not seem to be a clear pattern concerning the deletion of *hogy:* neither the verbs nor any of the subcorpora show characteristic distribution with respect to this option. We do not mean to draw the conclusion that there would not be any tendencies governing the choice of the speakers to use or drop

---

[7]   In the case of *mond,* we excluded those occurrences that were in the first person plural (*mondjuk* 'we say' or 'let us say'), as this functions as a discourse marker quite frequently.

[8]   These cover those items in which either the given verb had no subordinate clause of the relevant type, or in which there was some feature that generally blocks *hogy*-deletion, e.g., when the subordinate clause precedes the main clause.

the subordinator, only that these tendencies could not be captured through this search. The sample, as usually, could have been larger, and the query could have been conducted with more specified grammatical features (e.g., generating samples in which either all or none of the matrix sentences contain an overt pronoun associated with the subordinate clause), but it is quite likely that the differences between the speakers and random stylistic choices would raise difficulties even in that case. Therefore, when labeling the variants in phenomenon 8 (optional *hogy*-deletion), using "more frequent" to mark the variant with the subordinator is an arbitrary choice.
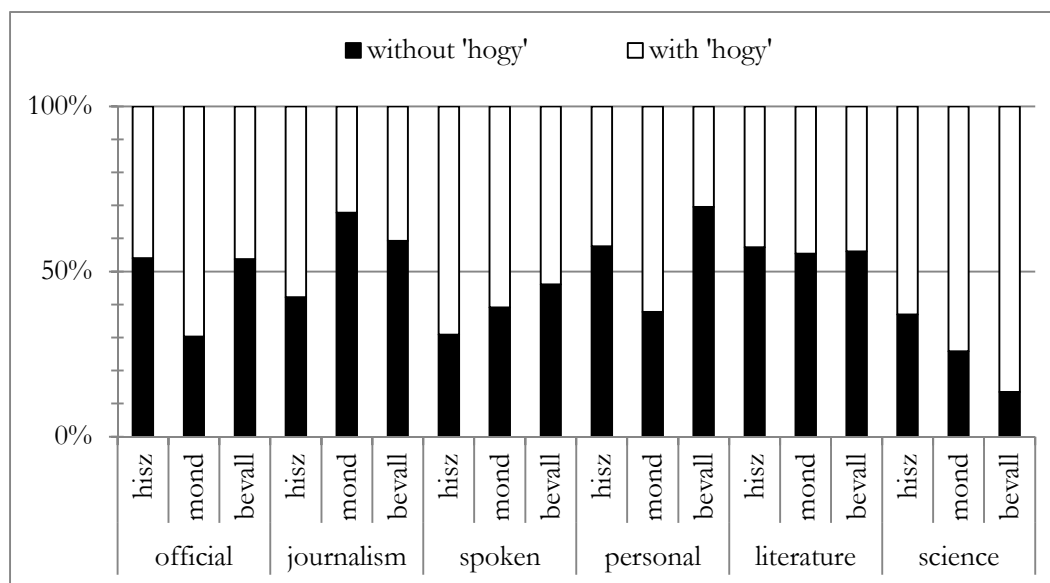

Figure 1: *The proportion of subordinate clauses with and without the subordinator* hogy

In the section of the test that focused on the use of *hogy,* another set contained sentences in which the verb of the main clause was *lehet* ('it may be'). In contrast with the verbs investigated above, the subordinate clause of this matrix verb would necessarily be headed by an overt complementizer (i.e., *hogy)* in MSH. Still, especially in non-formal registers, *hogy*-deletion seems to be spreading. The test contained five sentences of the type in example (5) below.[9]

(5) *Lehet,*   *(hogy)*   *én elmegyek   addig   fürdeni.*
    be.MOD.PRS.3SG (that)   I   go.PRS.1SG   till.then   bathe.INF
    'Perhaps I'll have a bath till then.'

The corpus search in this case ran with different settings for several reasons. As the deletion of the complementizer seems to be an innovative feature, we searched only those subcorpora (spoken, personal, journalism) that were thought to be more likely to contain instances of the innovative form (i.e., where the complementizer is deleted). Besides, the size of the random samples was meant to be bigger, containing a thousand items from each subcorpus. However, as *lehet* can occur in different structures, and many

---

[9] Many of the examples of *hogy*-deletion after *lehet* in our test stem from the collection of our colleague László Horváth, whom we sincerely thank for providing us with these data.

of these could not be filtered out automatically,[10] after the manual analysis there was only a smaller subset of the original set of three thousand sentences left that could be classified with respect to the presence or absence of the complementizer. Finally, we ran a search with the same settings in the Transsylvanian subcorpus, as we had an impression that this phenomenon could be more frequent in certain eastern dialects of Hungarian. Figure 2 summarizes the results of the queries.
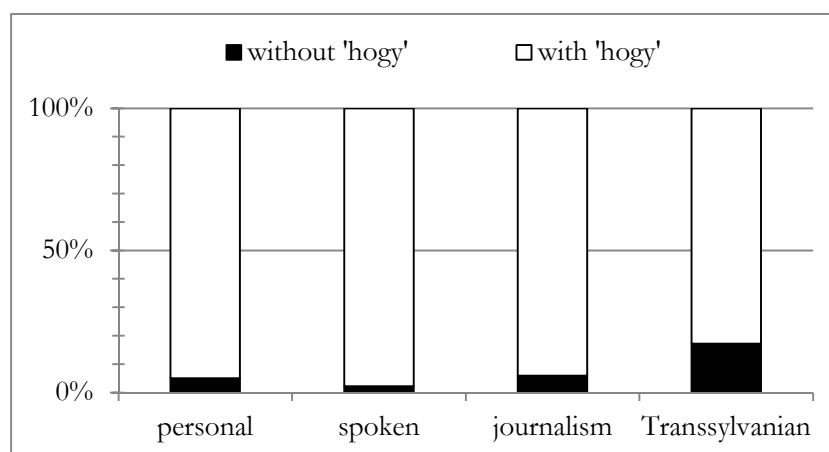

Figure 2: *Sentences with* lehet *'may be', with and without the subordinator* hogy

On the basis of the data (see Figure 2), one could draw the following conclusions. First, although the innovative, unmarked structure is rare, it could be attested in all these subcorpora. Second, the impression that this is a regionally rooted phenomenon is confirmed by the data, as the unmarked form occurred most frequently in the Transsylvanian subcorpus. Finally, and not surprisingly, the frequency of the conjunction-less subordinate sentences with *lehet* as a matrix verb is low compared to the three verbs examined above (those in the case of which *hogy*-deletion is optional in MSH).

Naturally, the question arises how to analyze the innovative form structurally: is it the case that a further matrix verb joined the ranks of those verbs that allow complementizer-deletion, or do the innovative structures require a different structural analysis? In the latter case, the interpretation could be that this is an instantiation of grammaticalization across clauses, and the original matrix verb developed a new function, surfacing as a sentence adverbial, with its original subordinate clause in its scope. This type of change (during which biclausal structures become reanalyzed as monoclausal ones) keeps recurring with *hogy*-clauses throughout the history of Hungarian (Haader 2001), and, as a result, original matrix verbs are reanalyzed either as sentence adverbials or as conjunctions. In fact, the same verb *lehet* with the adverbial *jól* 'well' gave rise to the complementizer *jóllehet* 'although' ('It can well be that X [proposition]' → 'Although X'). Besides, this type of change (clauses of the type 'it may be that' giving rise to 'maybe'-adverbials) is common cross-linguistically as well (Beijering 2010). As it is not possible to

---

[10] For instance, *lehet* can also function as a quasi-auxiliary, in which case the main verb is in the infinitive (e.g., *ezt meg lehet csinálni* 'this can be done'). These could be discarded automatically prior to generating the random samples by unselecting those cases in which there was an infinitive in the vicinity of the given verb. Still, numerous other types had to be removed during the manual analysis, as they were irrelevant from the point of the present analysis, like the frequent pattern in which *lehet* is a copula, occurring within a sentence containing a nominal or adjectival predicate (e.g., *okos lehet* '(s)he may be smart').

go into further details here (cf. Gugán 2015 for a more elaborate analysis), we can only suggest that many instances of the novel form are apparently ambiguous structurally (bi-clausal structure with omitted complementizer or monoclausal structure with a sentence adverbial), and it would be important to find the appropriate tests (potentially based on the prosodic pattern of the sentences) that could distinguish these two potential structures reliably.

The last set of sentences that focus on the presence or absence of *hogy* (phenomena 5, 6, and 7 in the test) do not contain matrix verbs at all. The striking feature of the given structure is that there is an adverbial that seems to govern a subordinate clause in this case, as in the examples (6)-(8).

(6) *Természetesen,*   *(hogy)*   *neki*   *van*   *igaza.*
    Naturally   (that)   (s)he.GEN   be.PRS.3SG truth.3SG
    'Naturally, (s)he is right.' [lit. '(S)he has got his right.']

(7) *Valószínűleg,*   *(hogy)*   *fogalma sincs*   *róla.*
    Perhaps   (that)   clue.3SG not.exist it.DEL.3SG
    'Perhaps (s)he does not even have a clue (of that).'

(8) *Nyilván,*   *(hogy)*   *az*   *én*   *ebédemet*   *ette*   *meg.*
    Obviously   (that)   the   my   lunch.1SG.ACC   eat.PST.3SG<3SG PV
    'Obviously, it was my lunch (s)he ate.'

This structure was first described in the seventies (see e.g., E. Abaffy 1976), but some highly sporadic instances can be attested already in Late Old Hungarian (Haader 2001). Owing to the truly unique nature of these structures, there were quite a few attempts to provide a structural analysis and/or an account of its development. A characteristic type of explanation is that this pattern is a result of the interference of two structures: an adjective that could subcategorize for a subordinate clause headed by *hogy* (9) and an adverbial that is the suffixed form of the same stem, but which could not have such a complement (10).

(9) *Természetes,*   *hogy*   *neki*   *van*   *igaza.*
     Natural   that   (s)he.DAT   be.PRS.3SG right.3SG
     'It's natural that he is right.'

(10) *Természetesen*   *neki*   *van*   *igaza.*
     Naturally   (s)he.DAT   be.PRS.3SG right.3SG
     'Naturally, (s)he is right.'

Among others, E. Abaffy (1976) takes this position, but she also points out that the reanalysis of the given adverbials as adjectives (a recurrent type of change with certain types of adverbials throughout the history of Hungarian) could also motivate the appearance of such structures. Nemesi (2000) surveys the available analyses, and votes for a type in which there are in fact two clauses, and the adverbial modifies a deleted matrix verb and a deleted expletive pronoun *(Valószínűleg igaz az, hogy* ['Probably ~~it is true~~ that'] → *Valószínűleg, hogy* ['Probably that']). Kenesei (2002) suggests that the functional element *hogy* in these cases is not an instantiation of the complementizer, and the structure itself is monoclausal, the adverbial being in the canonical position of sentence

adverbials. Finally, É. Kiss (2010) attributes the appearance of this structure to language contact with Rumanian. Her main argument is that this pattern is not quite compatible with the general structural characteristics of Hungarian, whereas several of the Romance languages use structures like this productively. Therefore, the structural description she offers relies on the structure hypothesized for the Rumanian pattern, meaning that this type of adverbials appears as the head of a special projection (SAP, Speech Act Projection) that subcategorizes for CPs. (For a more thorough review on the literature on this topic, see also the summary in É. Kiss 2010).

It is beyond the scope of the present paper to argue for any of these analyses. However, as a basis for comparing the test results, we also searched for the appearance of these structures in HNC. In this case, first we searched for all the instances three adverbials *(természetesen* 'naturally'*, valószínűleg* 'probably'*, nyilván* 'obviously'*).* Then we searched specifically for those cases when these items are followed by *hogy.* These data were then analysed per subcorpora, during which the irrelevant data were sorted out.[11]

Figure 3 shows the frequency of the adverbials occurring with *hogy* in each subcorpus. As the proportion of the adverbial with *hogy* is rather low, Figure 4 contains the occurrences of adverbial+*hogy* on a different scale, thus allowing the comparison of the subcorpora.
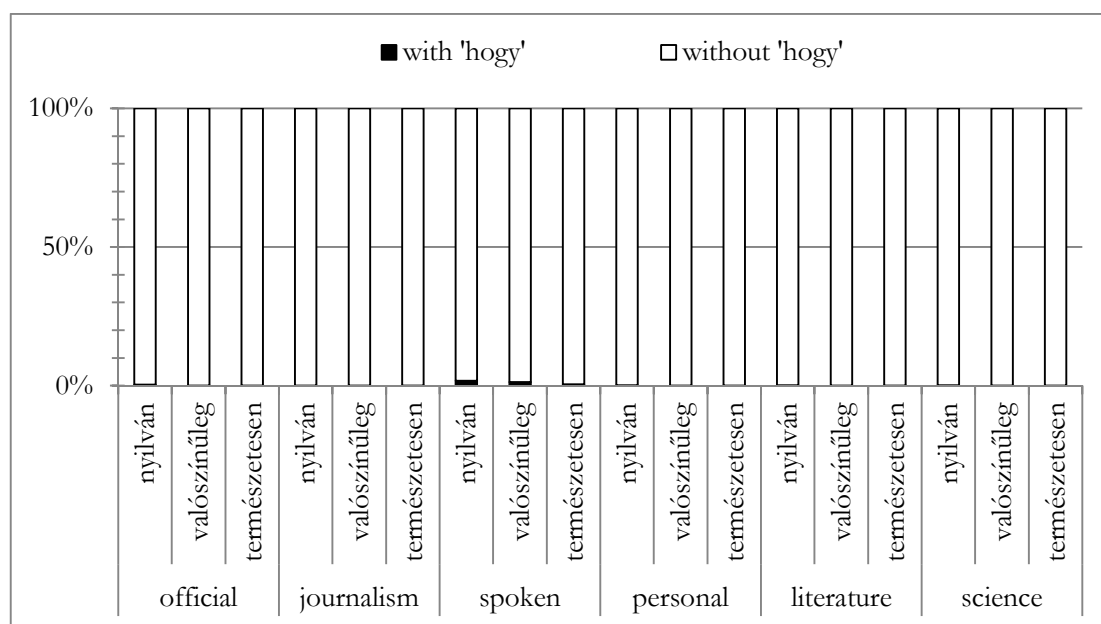

Figure 3: *Adverbials with and without* hogy *in the six subcorpora of HNC*

---

[11] I.e., those cases when the adverbial and the complementizer are adjacent, but the complementizer does not belong to the adverbial; this mostly meant hits in which the subordinate clause was an argument of the matrix verb, and the sentence adverbial modified that main clause as well, e.g.,

(i) *Úgy gondolom   természetesen,   hogy   […]*
     so think.1SG   naturally   that
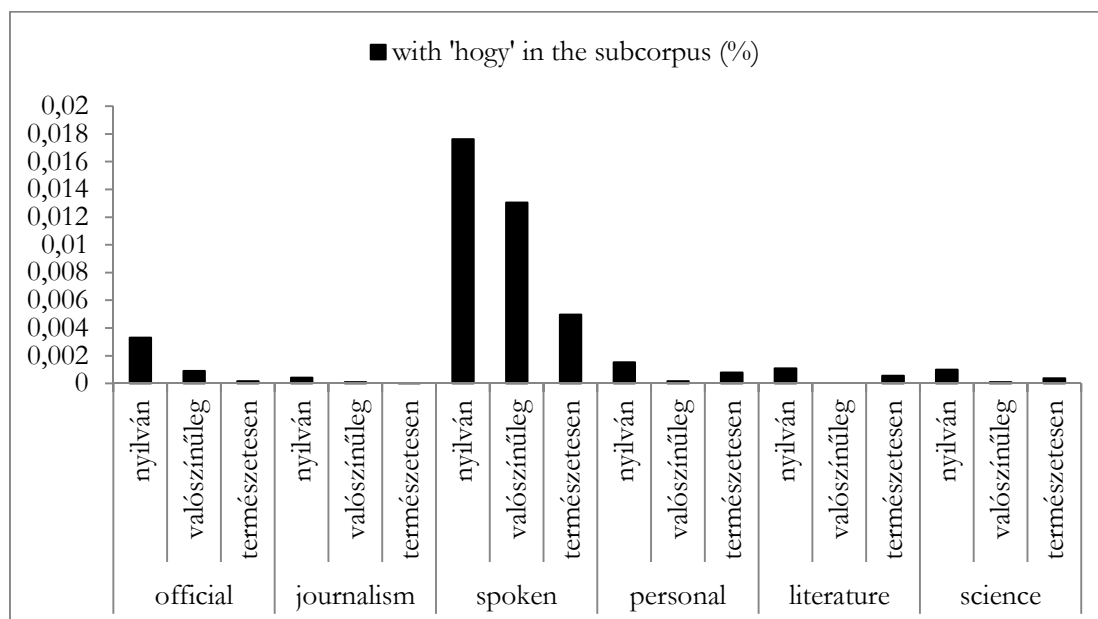   'Naturally, I think that […]'

Figure 4: *Adverbials with* hogy *in the six subcorpora of HNC*

The most obvious remark to be made is that the adverbial+complementizer structure is quite rare: even in the spoken subcorpus, out of all occurrences of *nyilván* 'obviously', less than 2% of them co-occurs with the complementizer *hogy*. Here the following characteristics of the subcorpora should be recalled. The subcorpus "spoken" (where the adverbial+complementizer structure appears to occur with some frequency) contains utterances representing both formal and informal registers; the subcorpus called "personal" contains texts from forum comments from the Internet displaying a more informal register; and the subcorpus "official" is a mixture of read aloud speech and spontaneous speech. Consequently, the data suggest that the appearance of the structure in question is characteristic of spontaneous speech (a speaking style in which the speaker has less time to plan and monitor his/her speech production), and not necessarily characteristic of domains that can be characterized as more informal, allowing for innovative and/or potentially stigmatized variants (contra Kontra 2003).

4. Unmarked object (phenomenon 9)

In MSH, all objects are marked with the accusative suffix *-t* (11a-b).

(11) a. *jön   a   hajó*
        come.PRS.3SG the   ship
        'the ship is coming'
     b. *látom   a   hajót*
        see.PRS.3SG<1SG   the   ship.ACC
        'I see the ship'

There is one regular exception: its use is not obligatory with nouns marked with a Sg1 or Sg2 possessive suffix (12a-b).

(12) a. *jön*   *a*   *fiam*
         come.PRS.3SG   the   son.1SG
         'my son is coming'
      b. *látom*   *a*   *fiam*
         see.PRS.3SG<1SG   the   son.1SG
         'I see my son'

As for the potential factors motivating the choice of the speakers to use the unmarked or the marked alternative, this seems to be totally optional grammatically, and neither of the forms is stigmatized or bound to registers. Diachronically, the unmarked accusative in this case is an archaic feature with parallels from e.g., Eastern Mansi (see Virtanen 2013); unfortunately, it is beyond the scope of this paper to survey the different historical explanations for this phenomenon (for details, see e.g., Korompay 1991, or É. Kiss 2014).

As there are no indications that register could be a factor concerning the choice between the two variants, the sample investigated from HNC was not filtered with respect to subcorpora.[12] We simply obtained random samples of nouns with Sg1 or Sg2 personal suffixes in the accusative, and manually classified the results according to the presence or absence of the accusative marker (besides, hits that were irrelevant for the present investigation were discarded). Figure 5 present the results of the corpus queries.
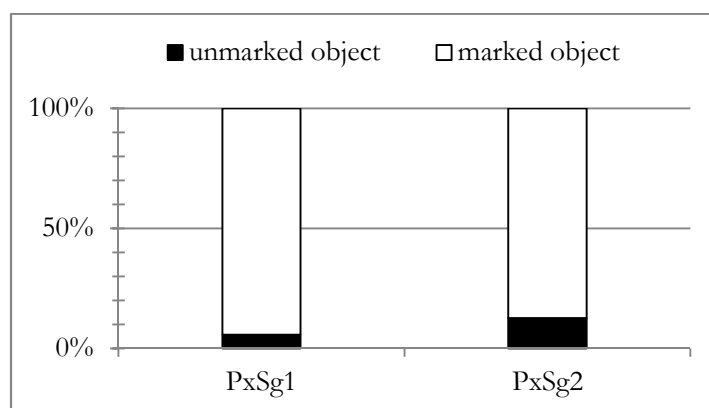

Figure 5: *Possessive-marked objects (with Sg1 or Sg2 personal suffixes)*
with and without the accusative marker

It has to be mentioned that unmarked objects are in fact more frequent in the corpus than the data above would suggest, as the investigation of similar samples with the nominative form (i.e., nouns with Sg1 or Sg2 personal suffixes in the nominative) revealed that unmarked accusatives are in many case wrongly classified and thus appear in the lists containing nominative forms as well.[13] However, even taking these findings

---

[12] The proportion of the two forms might still be different regionally, but this has not been investigated specifically.

[13] We also checked two samples of nouns with Sg1 or Sg2 personal suffixes in the nominative. It turned out that 18 out of the 1000 nouns bearing a first person possessive marker, and 105 out of the 1000 nouns bearing a second person possessive marker were in fact unmarked accusatives. Besides, it is also telling that in a recently conducted test in which participants had to fill in slots in sentences under time pressure (Kalivoda 2016), the majority chose the possessive marked variant without the

into account, the unmarked, archaic pattern seems to be less frequent, but absolutely not marginal.

5 Unmarked plurality of the possessor on the possessee (phenomenon 11)

The expression of possession with third person plural possessors also displays variation in MSH.

First, the possessor (in all person-number variations) can be either in the nominative or the dative (13a-b*).* As opposed to the optionality of markedness described above in the case of object marking, variation here is not arbitrary: the different case assignments are due to the different structures of these two patterns (for a detailed structural description of the possessive constructions, see Szabolcsi & Laczkó 1992).

(13)   a.   *Péter   háza*
           Peter.NOM   house.3SG
           'Peter's house'
    b.   *Péter-nek   a   háza*
           Peter-DAT   the   house.3SG
           'Peter's house'

Second, the choice of the suffix marking third person plural possessor on the possessee depends on the grammatical category of the possessor. If it is encoded as a pronoun, the marker of the Pl3 possessor on the possessee (14b*)* is different from that of the Sg3 possessor (14a), whereas the pronoun itself is not marked for plurality. However, if the possessor is encoded as a noun, then the possessee is unmarked with respect to the plurality of the possessor (and the possessor bears a plural marker, as in (15a-b).

(14)   a.   *az   ő   ház-a*
           the   he   house-3SG
           'his house'
    b.   *az   ő   ház-uk*
           the   he   house-3PL
           'their house'
(15)   a.   *a   fiú   ház-a*
           the   boy   house-3SG
           'the boy's house'
    b.   *a   fiú-k   ház-a*
           the   boy-PL   house-3SG
           'the boys' house'

Naturally, the possessor could be in the dative case in the above examples (14)-(15*)* as well. From the point of view of the present discussion, the relevant pattern is when the plural third person possessor is encoded as a noun in the dative as in (16a-b).

---

accusative marker in the task when the context rendered it likely that the object would be a possessee as well.

(16)   a. *a   fiú-k-nak   a   ház-a*
           the   boy-PL-DAT   the   house-3SG
           'the boys' house'
       b. *a   fiú-k-nak   a   ház-uk*
           the   boy-PL-DAT   the   house-3PL
           'the boys' house'

Ultimately, this variation can also be explained on a diachronic basis. The dative suffix on the possessor *(-nak)* is itself an innovation compared to the unmarked possessor (i.e., unmarked dative), albeit an ancient one, as it emerged during Late Proto-Hungarian. Originally, the possessor bore only the plural marker *-k*, and the possessee was either unmarked or marked with respect to pluraty (i.e., that of the possessor), so the pair that had coexisted first can be seen in (17a-b). The second member of this pair, i.e., the pattern showing number agreement (17b) is already obsolete in MSH.

(17)   a. *fiú-k   ház-a*
           boy-PL   house-3SG
           'the boys' house'
       b. *fiú-k   ház-uk*
           boy-PL   house-3PL
           'the boys' house'

There are good reasons to suppose that the non-agreeing form was the original pattern, and the two structures competed in Old Hungarian, but the non-agreeing form was prevalent in those texts that are closer to spoken language (Korompay 1991, 269). As for the dative-marked possessor in the same period, it was observed (Korompay 1992, 348) that whereas in structures where the possessee is not marked for the plurality of the possessor, the possessor can either be in the nominative or the dative *(fiú-k-0 ház-a* or *fiú-k-nak ház-a* were both frequently occurring types), in structures where the possessee is marked for plurality, the dative-marked form already seems to be obligatory *(fiú-k-nak ház-uk)*. Therefore, the variation that is characteristic of MSH (as in 15) is already present in Old Hungarian.

Naturally, the coexistence of these forms must have been characteristic of the period between Old and Modern Hungarian as well, which is illustrated by the fact that, according to the survey of the handbook of normative linguistics (Grétsy & Kovalovszky 1980, 350), representatives of the language reform movement (taking place between 1790-1820 approximately) propagated the exclusive use of the agreeing form both with the nominative and with the dative possessor, meaning that they must have been aware of the presence of variation in agreement. It is also instructive to see how the authors of the handbook of normative linguistics interpret this variation concerning MSH: they note that the use of the non-agreeing form is spreading in MSH, but they also recommend the use of the plural-marked possessee with dative-marked possessors, especially in that case when the structure appears in an intransitive sentence expressing possession as in (18).

(18)   *A   fiú-k-nak   van   ház-uk   /   (ház-a)*
        the   boy-PL-DAT   be.PRS.3SG   house-3PL   /   (house-3SG)
        'The boys have a house.'

Therefore, as in this case it cannot be excluded that the non-agreeing form is associated with a specific register (and this explains why the authors of the handbook find it less adequate in MSH), the data collected from HNC are arranged according to subcorpora, and in this case again we chose to investigate those three subcorpora that are thought to be closer representatives of informal language use. The query in this case was set for nouns in the plural dative form, with the verbs *van* 'there exists' or *nincs* 'there does not exist' in their immediate environment. As hoped, the search results contained many instances of sentences expressing possession ('He has got' / 'He does not have'), which is encoded in Hungarian with the possessor in the dative and the existential verb. (This way the data obtained through the corpus search can be easily compared to the test results, as the test also contained sentences expressing predicative possession). Figure 6 summarizes the results of this query. The random samples in this case contained 250 elements for each verb and each subcorpus, so altogether we analyzed 1,500 structures, but again, about a fifth of all data proved to be irrelevant from the point of view of the present discussion.
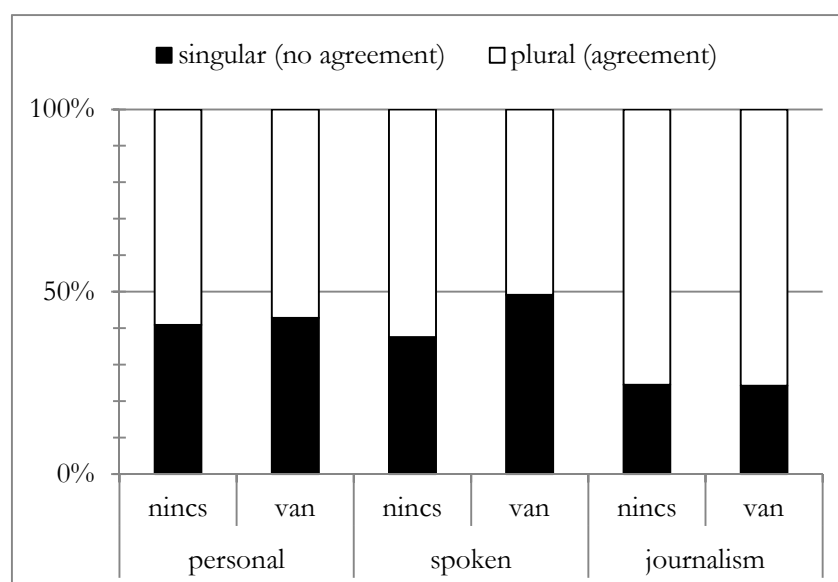

Figure 6: *Forms showing and not showing agreement in sentences expressing possession*

As it is apparent from Figure 6, although the agreeing form is in the majority with both verbs in all subcorpora, it could not be said that the non-agreeing form is infrequent; on the contrary, the speakers seem to use both forms widely.

6   Compounds of the type noun+verb (phenomenon 10)

There is a type of compounding in Hungarian in which the first stem of the compound is a noun, and the second is a verb, e.g., *apróhirdet* 'to post small ads', *bájcseveg* 'to do small talk', *agymos* 'to brainwash' (Kiefer 2000, 531). At first sight, these seem to consist of a verb and one of its arguments, the latter losing its appropriate morphological marker through the process of compounding, e.g., *város-t$_{ACC}$ néz* 'to go sightseeing', lit. 'watches the city'→ *városnéz*. However, this analysis would not be correct: these noun+verb type of compounds are backformations from pre-existing derivations in which the last stem is an action noun formed with the suffix *-ás/-és,* e.g., *apróhirdetés, bájcsevegés, agymosás.* Kiefer

(2000, 531) points out that sporadically, some compounds of this type may arise straight from a verbal phrase due to analogy, but this pattern of compounding is not (yet) productive in Hungarian. As these forms are morphologically transparent, all native speakers can interpret (and, theoretically, produce) the innovative, noun+verb compounds. However, the corpus query in this case could not help to survey the incidence of the novel forms, as such forms rarely occur in HNC: among the tested five compounds, only one *(ügyintéz* 'to administrate') was attested in the corpus as well. Therefore, this type of compound can by all means be assumed to be less frequent than the forms that can be the input of the word formation process (either the derived form with the action nominal, e.g., *ügyintézés* 'administration' or the noun + verb pattern, e.g., *ügyeket intéz*).[14]

Concluding this section, we have to emphasize that the aim of Experiment1 (see Section 3.2-3.3 in the article) was to single out linguistic variables that are subject to language change in MSH. For this purpose, we elicited grammaticality judgments concerning the above described variants, some of which we could have reliably labelled as "innovation" or "archaism" from a synchronic point of view prior to the grammaticality judgement test, and some of which we could only label as "less frequent" or "more frequent" on the basis of surveys in HNC. There were two reasons for collecting grammaticality judgments. First, we wanted to check whether any of the pairs that we referred to as "more" and "less" frequent corresponded to the opposition "conservative" and "innovative" from a synchronic point of view. Second, we wanted to check the validity of the previously assigned labels "conservative" and "innovative". It is important to note here, however, that we did not expect all the tested variants to fit into the innovation-archaism opposition. On the contrary, we included some phenomena which were thought not to be representatives of this opposition, that is, which were thought to show stable variation (e.g., phenomena 9 and 10, that is, optional *hogy*-deletion and unmarked accusative). We hoped that those cases that seemed to be prototypical instances of stable variation would serve as a basis for comparison, facilitating the selection of true innovations.

References

Bácskai-Atkári, Júlia and Éva Dékány. 2014. From non-finite to finite subordination: The history of embedded clauses. In Katalin É. Kiss (ed.), *The Evolution of Functional Left Peripheries in Hungarian Syntax*. 148–223. Oxford: Oxford University Press.

Beijering,   Karin.   2010.   *The   grammaticalization   of   Mainland   Scandinavian   MAYBE.* https://bells.uib.no/bells/article/view/39 (2012. 10. 12.)

E. Abaffy, Erzsébet. 1976. Valószínűleg, hogy…? *Magyar Nyelvőr* 100. 397–8.

É. Kiss, Katalin. 2010. Valószínűleg, hogy román kontaktushatás. In Katalin É. Kiss & Attila Hegedűs (eds.), *Nyelvelmélet és kontaktológia*. 223–238. Piliscsaba: PPKE BTK.

É. Kiss, Katalin. 2014. The evolution of functional left peripheries in the Hungarian sentence. In Katalin É. Kiss (ed.), *The Evolution of Functional Left Peripheries in Hungarian Syntax*. 9–55. Oxford: Oxford University Press.

---

[14]   A Google search in this case showed slightly different results: *tisztújít* [approximately: 'elect new leaders'] seem to occur with some frequency in the language of the press mainly, and there were also hits for *magánrendel* 'to do private consultations', *osztálykirándul* 'to go for a class outing' and *vészfékez* 'to push the emergency brake'. Naturally, this result is not surprising, as Google searches the entire web, therefore, new patterns can be found more easily.

Egedi, Barbara. 2014. The DP-cycle in Hungarian and the functional extension of the noun phrase. In Katalin É. Kiss (ed.), *The Evolution of Functional Left Peripheries in Hungarian Syntax*. 56–82. Oxford. Oxford University Press.

Givón, Talmy. 2001. *Syntax. An Introduction. Volume 1*. Amsterdam: John Benjamins.

Grétsy, László & Kovalovszky Miklós (eds.). 1980. *Nyelvművelő kézikönyv*. Budapest: Akadémiai Kiadó.

Gugán, Katalin. 2015. Hol volt? Hol nem volt? A tagmondattörlő grammatikalizációs folyamatokról a *lévén* és a *lehet* grammatikalizációja kapcsán. *Magyar Nyelv* 111. 38−53.

Haader, Lea. 1991. Az összetett mondatok. In Benkő, Loránd (ed.), *A magyar nyelv történeti nyelvtana I*. 728–741. Budapest: Akadémiai Kiadó.

Haader, Lea. 1995. Az összetett mondatok. In Benkő, Loránd (ed.), *A magyar nyelv történeti nyelvtana II/2*. 506–665. Budapest: Akadémiai Kiadó.

Haader, Lea. 2001. Mikrodiakrónia és változásvizsgálat. *Magyar Nyelvőr* 125. 354–371.

Harris, Alice C. & Lyle Campbell. 1995. *Historical syntax in cross linguistic perspective*. Cambridge: Cambridge University Press.

Heine, Bernd & Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

I. Gallasy, Magdolna. 1991. A névelő és a névelő-előzmény. In Benkő, Loránd (ed.), *A magyar nyelv történeti nyelvtana I*. 461–475. Budapest: Akadémiai Kiadó.

I. Gallasy, Magdolna. 1992. A névelők. In Benkő, Loránd (ed.*), A magyar nyelv történeti nyelvtana II/1*. 716–771. Budapest: Akadémiai Kiadó.

Kalivoda, Ágnes. 2016. *Szakít – félbe, szét, valakivel, valamit? Az igei komplexum feldolgozása produkciós feladatban*. Manuscript.

Kenesei, István. 1992. Az alárendelő mondatok szerkezete. In Kiefer, Ferenc (ed), *Strukturális magyar nyelvtan 1. Mondattan*. 529–714. Budapest: Akadémiai Kiadó.

Kenesei, István. 2002. Hányféle igazság van? *Magyar Nyelv* 98. 39–49.

Kiefer, Ferenc. 2000. A szóösszetétel. In: Kiefer, Ferenc (ed), *Strukturális magyar nyelvtan 3. Morfológia*. 519–568. Budapest: Akadémiai Kiadó.

Kontra Miklós (ed.) 2003. *Nyelv és társadalom a rendszerváltáskori Magyarországon*. Budapest: Osiris.

Korompay, Klára. 1991. A névszójelezés. A névszóragozás. In Benkő, Loránd (ed.*), A magyar nyelv történeti nyelvtana I*. 259–318. Budapest: Akadémiai Kiadó.

Korompay, Klára. 1992. A névszójelezés. A névszóragozás. In Benkő, Loránd (ed.), *A magyar nyelv történeti nyelvtana II/1*. 321–410. Budapest: Akadémiai Kiadó.

Nádasdy, Ádám. 2008. *Prédikál és szónokol*. Budapest: Magvető.

Nemesi, Attila László. 2000. A *természetesen, hogy…* típusú szintaktikai szerkezetről. *Magyar Nyelvőr* 124. 430–442.

Oravecz, Csaba, Tamás, Váradi & Bálint, Sass. 2014. The Hungarian Gigaword Corpus. In *Proceedings of LREC 2014*.
http://www.lrec-conf.org/proceedings/lrec2014/pdf/681_Paper.pdf

Szabolcsi, Anna & Tibor Laczkó. 1992. A főnévi csoport szerkezete. In Kiefer, Ferenc (ed.), *Strukturális magyar nyelvtan 1*. Mondattan. 179–298. Budapest: Akadémiai Kiadó.

Virtanen, Susanna. 2013. Contextual function of noun marking in the direct object–marking system of Eastern Mansi. In Csepregi, Márta, Kubínyi Kata & Jari Sivonen (eds), *Grammatika és kontextus: új szempontok az uráli nyelvek kutatásában III*. Uralisztikai Tanulmányok 20. 305–317. Budapest: ELTE Finnugor Tanszék.

Towards capturing implicit innovative language attitude using an auditory Implicit Association Test[*]

Andrea Deme, Katalin Gugán, Bálint Sass, Katalin Mády

Since the birth of sociolinguistics the localization of innovative speakers has been regarded as a key issue in the study of language change. For this purpose, researchers traditionally categorize the speakers of a speech community on the basis of demographic and socioeconomic features; however, these parameters prove not to be sufficient to identify innovative speakers in all cases. It may be argued, however, that the speaker's implicit attitude towards linguistic innovations may also be captured and may be a good indicator of the speaker's innovative linguistic behavior. This line of research is not yet well elaborated on, probably due to the complexity of the attitude construct (which makes the measurement of implicit attitudes a challenging task), and the difficulty of grasping attitude towards linguistic innovations as such. The present study aims at addressing the potentials present in this aspect of sociolinguistic investigation. We review the psychological literature on the attitude construct and propose that a method borrowed from social psychology, the Implicit Association Test (IAT), may be adapted for the measurement of implicit attitude towards linguistic innovations if used with a linguistic variable that is subject to an ongoing language change as the test variable (or target). We report a pre-test conducted for variable selection for the adaptation of the IAT, and analyze data gathered by means of this newly created method for capturing implicit innovative linguistic attitude. We propose that this new IAT may be a useful tool in language change studies.

Keywords: IAT, implicit language attitude, language attitude, innovative attitude, language change, sociolinguistics

1   Introduction

Traditionally gender, age and socioeconomic status are regarded as important factors of language change, as they are suggested to enable to us to designate innovative speakers, the key figures of linguistic change (see. e.g., Labov 1980). These factors, however, do not appear to be sufficient in the identification of innovative speakers in all cases (see also Labov 1980). As a possible solution to this problem we argue that new methods in the detection of innovative speakers are necessary which allow researchers to identify potentially innovative linguistic behavior.

Language attitude, in particular implicit language attitude is another factor that is often addressed in sociolinguistics, but mostly in studies where the issue of stereotypes and language-based prejudice is addressed through the investigation of implicit attitudes towards linguistic variation (e.g., towards dialectal or accented speech) (see e.g., Pantos 2010 and Pantos & Perkins 2013 amongst others). Accordingly (and also based on several other sociologically important reasons), there is a constantly growing body of research investigating implicit language attitudes, where language attitude is generally defined as a disposition that is evoked by language, because speech is assumed to provide cues based on which a listener may assign supposed group memberships to the speaker (see e.g., Preston 2003). In a relatively smaller amount of research, attitude is also

considered as a disposition towards linguistic variants, like, for instance, linguistic innovations, and according to these interpretations, attitude may also play an important role in language change (see e.g., Hopkins 1977). Yet, when it comes to this possible interpretation of attitude and what it may offer the investigation of language change, only the surface has been scratched.

In the present study we propose a solution to the problem of locating innovative speakers by expanding the most common (or perhaps the traditional) interpretation of language attitude, and by proposing that innovative speakers may be characterized by a specific implicit attitude towards linguistic innovations as such. In this paper we also propose a possible way of operationalizing this specific implicit attitude for linguistic research.

We argue that the implicit attitude a language user may hold towards language change and linguistic innovations may be a crucial component in the user's involvement in language change processes and thus it should also be considered in studies of language change. To the authors' knowledge, this interpretation of implicit attitude is so far rarely studied empirically, probably for the following two reasons. The first may be the duality of attitudes and the nature of the implicit component (i.e., the nature of *implicit attitude*) which makes attitude a very difficult phenomenon to examine. The second may be the difficulty of grasping attitude towards linguistic innovations as such, since linguistic attitudes are generally considered only with respect to socially well-defined speaker groups and not towards more abstract linguistic structures or concepts; we discuss both of these issues in more detail in the following sections. In the present study we argue that a specific aspect of implicit innovative language attitude may be captured by measuring the automatic (implicit) evaluative reactions to innovative linguistic forms, and we propose a possible solution for measuring this specific aspect of implicit attitude by adapting the Implicit Association Test (IAT), a measurement tool of social psychology. We also claim that by applying this test for measuring implicit attitudes, we also overcome the difficulty of grasping abstract linguistic categories (as innovative or conservative linguistic variants) for testing, as already demonstrated by previous research.

In accordance with the above, the aim of the present paper is to elaborate on the proposal of interpreting implicit attitudes as implicit attitudes towards linguistic innovations, and to report on the process of creating the IAT capable of capturing implicit attitude of language users towards Linguistic Innovations (this test will hereafter be referred to as LI-IAT). We argue that for this purpose, a linguistic variable that is subject to an ongoing language change must be identified and used in an auditory IAT paradigm. We suggest that through capturing this specific aspect of implicit attitude towards linguistic change, the LI-IAT may be a useful tool in the study of language change in the future: as the LI-IAT measure can be an indicator of implicit attitudes towards language change, it may enable us to identify potentially innovative language users and thus to investigate ongoing language change processes reliably. It seems to be appropriate to also anticipate the long-term objective of the authors: our goal is to map the specific aspect of implicit linguistic attitude captured by LI-IAT to several types of language change, e.g., sound change, morphological change, or syntactic change. This way, we plan to assess which types of language change the LI-IAT measure can be a reliable indicator of. The study presented here is the first step in this process.

The present paper is structured as follows. Section 2 outlines the psychological background of the attitude construct for the study of attitudes in sociolinguistics. We provide a brief summary of the interpretation of language attitude in sociolinguistics while also offering a new interpretation of innovative language attitude. After reviewing

the relatively small number of previous studies on implicit innovative language attitude, we conclude by formulating our aims and discussing how all the above lead to the implementation of the auditory IAT in the measurement of implicit attitude towards linguistic innovations. Section 3 and 4 present two experiments: the first is conducted for variable selection for the adaptation of the IAT, while in the second some preliminary data are analyzed which were gathered by means of this newly created LI-IAT for capturing implicit innovative linguistic attitude. Section 5 provides a brief overview and discusses possible applications of the LI-IAT in language change studies, as well as future work.

2   Background

2.1   The attitude construct

The attitude construct has continuously been a topic of interest in the psychological literature resulting in complex and to some extent even diverging definitions of the concept. However, it is beyond the possibilities of the present study to review and reflect on this diversity of research and theory; we can only venture to give a brief and basic (thus necessarily simplifying) description of it to provide the theoretical basis of the present study.

According to the most popular model, the *expectancy-value model*, attitude is a summary of *evaluation*, where the evaluative meaning arises inevitably and spontaneously as a result of cognitive processes, namely *associations*. In this model attitude can be captured as an association between the attitude object and valued *attributes*: notions that are eligible to express *valence* (i.e. evaluative meaning),[1] such as the dimension of *good–bad*, *pleasant–unpleasant*, or *harmful–beneficial* (Ajzen 2001). Attitude, namely the evaluation-based categorization, or the measure of favorability is to be differentiated from *evaluation-free categorization* or *sorting* (for instance sorting food made of vegetables and food made of non-vegetable ingredients; for evidence of this differentiation, see e.g. Cacioppo et al. 1996 with respect to the field of neuropsychology), and from the notion of *affect*, which has been described as "states that contain degrees of valence as well as arousal" (Ajzen 2001: 29). In some theories affect is even assigned precedence to over cognition, i.e. attitude formation (see the *affective primacy hypothesis* in Ajzen 2001 and its references).

2.2   The model of dual attitudes

In the literature of psychology there is general agreement that the evaluation of psychological objects is inevitable and spontaneous. However, attitude is not necessarily univalent, that is, many circumstances may facilitate the development and holding of more than one attitude towards the same attitude object. According to one of the most influential theories, the *model of dual attitudes*, the duality of attitudes lies within attitude change: when attitudes change, the new attitude may not fully replace the older, more habitual attitude, but the two may keep co-existing which results in dual attitudes, i.e., in two simultaneous but not necessarily congruent evaluative reactions to the same object (Wilson et al. 2000). For instance, early acquired (and even repressed) prejudice may co-

---

[1] Valence in psychology is defined as the intrinsic attractiveness (positive valence) or aversiveness (negative valence) of an event, object or situation (see Frijda 1986, 207).

exist with later created egalitarian views. According to the theory of Wilson and his colleagues this duality may be grasped by differentiating between *explicit* and *implicit* attitudes. *Implicit attitude* is an evaluative disposition considered to be subconscious (outside of conscious awareness) (see also Greenwald & Banaji 1995), it is habitual, it is based on unidentified (or inaccurately identified) traces of past experience (Greenwald & Banaji 1995) and environmental impact (Karpinski & Hilton 2001), and it shapes the interaction with the attitude objects remarkably (or at least influences implicit or uncontrollable responses that one might not make an attempt to control) (Greenwald & Banaji 1995, Wilson et al. 2000). *Explicit attitude*, on the other hand, is considered to be more recently constructed, deliberately formed, or in other words, consciously accessible. Consequently, explicit attitude is the disposition we can report on directly. This is the case when we answer questions about our preferences, for instance. In such a context, implicit attitudes are "unavailable". Additionally, it should be emphasized again that implicit and explicit attitudes may diverge. Therefore, data gathered through explicit evaluative questions will necessarily be able to reflect consciously available, i.e., explicit attitudes exclusively, while implicit attitudes which may have the opposite valence of explicit attitude are almost always left unrevealed by direct questions and questionnaires (see e.g., Horwitz & Dovido 2015 on diverging explicit and implicit attitudes towards wealthy people or the results of Pantos 2010 and Pantos & Perkins 2013 on diverging explicit and implicit attitudes towards accented speech).

Before we further elaborate on the effect of the duality of attitudes on attitude measurements (and on other factors that may also have an impact), we briefly discuss another aspect of attitude, namely *ambivalence,* that should clearly be differentiated from *duality* described above. Ambivalence is the co-existing positive and negative disposition toward the same object, but in this case, the conflict does not stem from the different "layers" of attitude (that differ in conscious availability), but the evaluation of the same objects on *different dimensions* resulting in a conflict within the cognitive component or in a conflict between affect and cognition (see e.g., Ajzen 2001). We present one example that illustrates this differentiation. In one of their studies MacDonald and Zanna (1998) asked male participants to evaluate feminist candidates in a job interview situation on two dimensions: *admiration* and *affection.* According to their results, males rated feminists positively on the dimension of admiration, but negatively on the dimension of affection, that is, they were proved to be holding ambivalent attitudes toward feminists, which was not the result of the implicit–explicit opposition, since evaluation on both dimensions was assessed on the basis of self-report. Therefore, the concept of ambivalence is clearly out of the scope of the present discussion, and will not be further discussed in the present study.

As already mentioned, the duality of attitudes, in other words, the separation of implicit and explicit attitudes poses a problem to attitude measurements, as data gathered through explicit evaluative questions will necessarily be able to reflect explicit attitudes exclusively. However, in many cases (e.g., in the case of the evaluation of linguistic innovations, i.e., innovative linguistic behaviour) we may assume that it is rather the implicit attitudes that are of interest. Moreover, the issue of the most frequently studied response bias, "socially desirable responding" (Paulhus 1991, 17), also comes into play when one investigates attitudes towards a socially sensitive area, at least if one does so through direct questions, self-reports or other explicit measurements. According to studies in social psychology, in those cases, when the attitude object is considered to be a socially sensitive object, attitudes measured by explicit evaluative questions tend to show a social desirability bias, i.e., in responding the informants try to respond according to

their beliefs about what a socially more acceptable response is and not in a way that reveals how they actually feel or believe (Holtgraves 2004). To give a basic impression of what this statement means, we cite Holtgraves' examples: according to studies conducted in the USA, people tend to overreport their engagement in socially desirable behaviors, such as attending religious services and voting, but underreport engaging in socially undesirable behaviors such as substance abuse (see Holtgraves 2004). According to Holtgraves (2004), social desirability operates as an "editing process": participants retrieve the requested information (e.g. the answer to an explicit question), but they also evaluate it before responding: they assess whether the response would make them look good or not. If truthfulness interferes with social desirability, respondents may respond according to the latter. Generally speaking, both the duality of attitudes and the social desirability bias pose a serious problem to the investigation of implicit and even to explicit attitudes. As direct questions may only reflect explicit attitudes, attitudes in connection with stereotypes and prejudice are even more difficult to reveal, due to the fact that informants may often tend to respond according to socially more acceptable attitudes (i.e., according to social solidarity and equality), and the attitudes that are not in line with these ideal dispositions may remain "covered".

Arriving at the focus of our present study, the consequences of the above factors can be summarized as follows. The measurement of explicit and implicit attitudes towards attitude objects, in our case, specific (innovative) language forms, is a demanding task, but it is of great importance. In certain areas of behavior, including language behavior or language use, some attitudes are a matter of prestige, while others are incorrect, substandard, not appropriate or stigmatized. Therefore, when investigating these areas the issue of socially desirable responding should also be taken into account. Moreover, in the case of innovative linguistic behavior, i.e., in the acceptance or use of innovative language elements it may well be assumed that implicit attitude is a key component. (Probably almost everyone had the experience of a friend who has expressed some negative opinion about a stigmatized linguistic form, and used it him- or herself just a few minutes later.). Implicit attitudes are, however, out of conscious awareness. These problems are serious, but can be overcome by using implicit attitude measures to detect innovative linguistic attitude.

2.3   Measurements of implicit attitudes: the Implicit Association Test

To gain access to implicit cognitions (a domain not reached by self-report measures) several solutions have been tested (for a short summary, see e.g. Karpinski & Hilton 2001). Among them, one of the most influential and widely used techniques is the Implicit Association Test (IAT). The IAT is a simple evaluation-free categorization task where implicit preference, i.e., positive attitude towards the attitude object is only deduced from response latency (i.e., reaction time, RT). The IAT is based on the notion of attitude being an association between the attitude object (*target*) and attributes with positive and negative valence (*attribute*). The principle of the method is that the ability to quickly sort the target items and attribute items to their corresponding categories reflects the strength of association within the two pairs of opposing target–attribute categories that have to be sorted together. Accordingly, in the IAT paradigm RT data are considered to be correlates of attitude strength (see e.g. Greenwald et al. 1998).

Figure 1 illustrates the basic IAT design, and in the next paragraph we will briefly go through the illustration to demonstrate how the IAT works in practice.

	Categories (and category labels)	Items
Target	BIRD	*cardinal, warbler, blackbird, robin*
	INSECT	*cicada, locust, bee, mosquito*
Attribute	PLEASANT	*cuddle, happy, smile, joy*
	UNPLEASANT	*abuse, crash, disaster, grief*

	Left	Right
Block 1:	BIRD	INSECT
Block 2:	PLEASANT	UNPLEASANT
Block 3:	BIRD OR PLEASANT	INSECT OR UNPLEASANT
Block 4:	BIRD OR PLEASANT	INSECT OR UNPLEASANT
Block 5:	INSECT	BIRD
Block 6:	INSECT OR PLEASANT	BIRD OR UNPLEASANT
Block 7:	INSECT OR PLEASANT	BIRD OR UNPLEASANT

Figure 1: *Illustration of the Implicit Association Test (IAT), adapted from Vande Kamp (2002, 3)*

Before the test, participants familiarize themselves with the *categories* and the corresponding *items* to be used in the test. There are always two opposing target categories (here, BIRD vs. INSECT) and two opposing attribute categories (here, PLEASANT vs. UNPLEASANT), resulting in a total of four categories and four category labels. In Block 1 participants acquire the assignment of the left and right sides (of the computer screen) to the two *target* categories (BIRD vs. INSECT) by sorting the target items appearing in the middle of the screen, according to the target labels that are visible in the two upper corners (sorting administered by key press, usually E for left, and I for right). Next, in Block 2 the participants also learn the assignment of the two *attribute* categories (PLEASANT vs. UNPLEASANT) by sorting the corresponding attribute items according to the two attribute category labels visible in the upper corners again. In Block 3 the attribute and target category labels appear *together* with the conjunction *or* (e.g., BIRD OR PLEASANT on the left and INSECT OR UNPLEASANT on the right), and attribute and target items are to be sorted simultaneously: target items (e.g., *cardinal*) according to target labels, and attribute items (e.g., *happy*) according to attribute labels (note that although target and attribute items appear in succession, participants still sort target items according to target labels, and not attribute labels, thus no explicit evaluation is required). This task is repeated in Block 4. In Block 5 the inverse assignment of target labels is acquired by sorting target items again, according to the new (inverse) set-up of labels. In Block 6 attribute and target labels appear together again in the second attribute + target combination (since attribute labels are displayed unchanged), and participants sort all the target items and attribute items (similarly to Block 3 and 4). In Block 7 the task of Block 6 is repeated.

The IAT is based on the idea that if highly associated concepts share the same side (and the same key response), participants are able to categorize items much faster than in the opposite combination (when weakly associated concept share sides and key responses). Therefore, to calculate the IAT effect the *congruent* (i.e., expected stronger association, e.g., *bird* and *pleasant*) and *incongruent* (i.e., expected weaker association, e.g., *insect* and *pleasant*) blocks must be compared. As one attribute + target alignment is

predicted to be easier than the other attribute + target combination, Block 4, 5, 6 and 7 basically consists of "congruent" or 'easy' and "incongruent" or 'hard' blocks. The IAT effect is quantified for each participant as a so called *D measure*: trials greater than 10,000 ms are deleted; "inclusive" (pooled) standard deviation for congruent blocks (Block 3 and 6, in the first setting), and incongruent blocks (Block 4 and 7) are computed; the mean latency of the congruent block is subtracted from the mean latency of the incongruent block ($Mean_{Block\ 6} - Mean_{Block\ 3}$ and $Mean_{Block\ 7} - Mean_{Block\ 4}$); each difference is divided by the corresponding pooled (inclusive) standard deviation; *D* equals the equal-weight average of the two resulting ratios (for further description see Greenwald et al. 2003). The value of *D* is normalized between −2 and 2 where effect size criteria meet the requirements of (and thus are interpreted similarly to) Cohen's *d:* $-.15 < D < .15$: *no effect;* $.15 \leq D < .35$ or $-.15 \geq D > -.35$: *weak effect;* $.35 \leq D < .65$ or $-.65 \geq D > -.65$: *medium effect;* $.65 \leq D$ or $-.65 \geq D$: *strong effect*, where positive values represent implicit bias in the *congruent* direction, and negative values reflect implicit bias in the *incongruent* direction (see e.g., Greenwald et al. 2003).

IAT was originally designed for sorting visual stimuli (strings or pictures). This is because the majority of implicit attitude detection studies is concerned with the issues of prejudice detection and stereotypes (i.e., attitudes toward racial and ethnical minorities, overweight people, etc.), and stereotypical groups can easily be represented with typical names or faces. However, there is no doubt that in human interaction socially meaningful variation of linguistic forms is also a common way to identify group membership, thus activation of stereotypes triggered by linguistic forms is a key issue in sociolinguistics and social psychology and should not be neglected either (Campbell-Kilber 2012). Recognizing the above, the IAT paradigm was also tested (and tested successfully) with auditory prompts. In his dissertation, Vande Kamp (2002) demonstrated that the auditory IAT is a reliable and useful extension of the IAT that consists entirely of visual stimuli, either if the sounds to be sorted are words (as given in Figure 1), or if the speaking voices are socially meaningful (e.g., the voice of European-American or African-American speakers).

In some studies IAT was also used to address questions that are relevant particularly from a sociolinguistic perspective. Pantos (2010) and Pantos and Perkins (2013) investigated foreign-accented speech versus non-accented speech. Their findings demonstrated implicit bias toward the non-accented speech, that is, they provided evidence that speakers favor their native language and their own accent over foreign-accented speech. Furthermore, the authors also obtained explicit measures (by means of self-report) which indicated an explicit bias in the opposite direction, that is, explicitly informants seemed to favor the foreign-accented speaker. These findings demonstrate that implicit and explicit attitudes are separable and often also diverging constructs, and they also exemplify how the so called social desirability bias may exert its influence on linguistic attitudes.

Using the auditory IAT paradigm, Kathryn Campbell-Kibler (2012) provided further insights. She recognized that studying implicit associations in the research of attitudes should not necessarily be restricted to the investigation of attitude object–value alignment (as in the case of the accented–non-accented speech and good–bad category, for instance), but may also be used to investigate any kind of alignments that may be meaningful sociolinguistically. She claimed that sociolinguistics may also benefit from studying association strength between linguistic forms (which activate stereotypical groups) and stereotypically associated concepts (region, education, socioeconomic status, etc.). Campbell-Kibler used auditory language variables (e.g., *talkin'* vs. *talking*) to activate

stereotypes and associated them with professions and the names of prototypical northern and southern states to gather some insight into the *meaning* of previously detected stereotypes (reported by sociolinguistics in many previous studies). Campbell-Kibler created several IATs using visual and auditory stimuli, and she used abstract metalinguistic category labels referring to linguistic variants as e.g., –IN and –ING (referring to the *talkin'* versus *talking* opposition). Both of these solutions proved to be appropriate for the categorization test the IAT is based on, and demonstrated that the difficulty of grasping abstract linguistic categories may be overcome by using the IAT paradigm and simple metalinguistic labels.

To summarize shortly, the above studies demonstrated that i. IAT is a useful tool in implicit attitude detection if used with auditory prompts, ii. IAT may be used to detect sociolinguistically relevant attitudes, and iii. that the use of IAT should not be restricted to the detection of stereotypes.

2.4 Attitude in linguistic studies and the interpretation of the concept proposed in the present study

In sociolinguistics the term *language attitude* is predominantly used for evaluative disposition towards linguistic variants in language variation that may serve as cues for the listener to identify the speaker's group membership (see e.g., Preston 2003). This is particularly true in the case of Hungarian sociolinguistics where language attitude is used and studied exclusively with respect to regional variants of Hungarian, or (to a lesser extent) foreign-accented Hungarian speech (see e.g., Kiss 2000 and references therein). This is most probably due to the fact that the study of language-based evaluation in sociolinguistic research is in many cases motivated by prejudice detection and detection of stereotypes associated with stereotypical groups, and activated by typical language use of the group members (e.g., accents or dialects).

In Hungarian research, most of the studies used explicit attitude measures to draw conclusions on attitude in general: they elicited the set of beliefs on language norms and standards combined with beliefs about socially expected responses to the assumed substandard forms. Additionally, some studies also introduced techniques to elicit a manifestation of implicit attitude. Sándor et al. (1998), for instance, used the *matched guise paradigm* (see Lambert et al. 1960, and e.g., Mac-Farlane & Stuart-Smith 2012 for a later adaptation of the technique in the study of sound change processes). In this paradigm the participants are told that they would hear several different speakers, and their task is to evaluate these speakers on the basis of the (recorded or live) utterances on several dimensions, e.g., intellect or reliability which are considered to be implicit inquiries about the likeability of the person in the analysis. In reality, however, the instruction introduces a deception, since all of the utterances attributed to different speakers are provided by only one speaker varying his/her own production by shifting styles between utterances. As a result, the matched-guise paradigm has two advantages because of which it qualifies as a more sensible choice to measure attitudes than direct-questions or questionnaires. First, the listeners actually evaluate linguistic variants and not speakers on the given dimensions, since voice-specific effects which would cloud the relationship between the linguistic variable and the evaluative response are controlled for (i.e., the only difference between the utterances is the linguistic variable under consideration). Second, the listeners evaluate linguistic variants through answering implicit questions, since they are not asked to express favorability, but to assign cognitive, physical or other capabilities or characteristics to the speaker. In the study of Sándor and colleagues (1998) one group of

listeners heard a speaker speaking only standard Hungarian, while another group heard him speaking only his native regional dialect of Hungarian. As expected on the basis of previous literature on language attitude and prestige forms, the speaker was preferred (assessed on the basis of evaluation on personality dimensions) when he used standard Hungarian.

It is important to note, however, that implicit methods, such as the matched-guise paradigm do not necessarily (or may not at all) reflect implicit attitudes, as one might always have the chance to catch his or her own stereotype activating while forming a reaction to an implicit question, and thus the social desirability bias may arise as an editing phase before giving a response to the question (as proposed by Holtgraves 2004). In order to detect implicit attitudes (and to overcome the limitation of self-reports), psychological methods may provide the optimal solution, as they do not require the participants to report on a subjective assessment, but infer attitude from other measures. One of these and probably also one of the most influential ones is, as already mentioned, the IAT paradigm.[2] In sociolinguistic studies, however, these psychological methods are used only to a limited extent, at least with respect to the issues addressed by these implicit methods. To the authors' knowledge, most of these applications involve the IAT (or a modification of it), and study only stereotypes or prejudice against linguistic variables that represent a stereotypical group of people (see e.g., the studies already mentioned: Pantos 2010, Pantos & Perkins 2013 investigating attitudes towards accented speech). Only a few studies ventured to further explore the potentials of these methods in sociolinguistics, like e.g., Campbell-Kibler who investigated sociolinguistic meaning by aligning the linguistic variables with concepts that did not "simply" express valence in an IAT (discussed above), but may index some other sociolinguistically relevant meanings.

Returning to the interpretation of the concept of language attitude in sociolinguistic studies, we argue that besides referring to the disposition towards the language use of stereotypical groups, this term's use can be broadened to encompass attitude towards language itself (as indicated also by e.g., Hopkins 1977). More specifically, language attitude may also be interpreted as a disposition towards linguistic phenomena, for instance, linguistic innovations. In this interpretation, innovative language attitude is conceived as a disposition that is held against innovative linguistic variants (or even towards a group of innovative linguistic variants), and in this way it is interpreted as an attitude towards language change itself. But what would we gain if we interpreted language attitude as an attitude held against linguistic variants (and not towards a group of people they might represent if this group is existent at all), and innovative language attitude as an attitude towards linguistic innovations and language change?

As Labov pointed out (emphasized in the review of Milroy and Milroy 1985, referring to Labov 1980), there are certain groups of the society who use language innovations more than others, and these are the groups who basically initiate the diffusion of linguistic innovations (see e.g., Milroy and Milroy 1985). Hence, Labov (and a great number of scholars after him) concluded that the localization of these groups, i.e., the groups of innovative speakers is a key issue in the study of language change. However, Labov also demonstrated in his studies that these groups cannot be localized easily, as these groups are not easy to characterize only by means of some demographic features or the speakers' socioeconomic status (see also the comments from Milroy &

---

[2]   There are, of course, many other implicit methods, like the Go/No-Go Association Task, priming tasks, etc., for a summary see e.g., Gawronksi & de Houwer (2014).

Milroy 1985). Luckily, demographic mapping is not necessarily the only solution to identify innovative speakers, if we assume that these speakers are characterized by a specific kind of implicit attitude towards linguistic innovations.

Based on what is known about attitude from social psychology, automatic preference and usage of certain linguistic forms can be considered as the manifestation of implicit attitude. Therefore, we argue that if we are able to identify some innovative forms in a given language (that are currently subject to an ongoing language change), and we are also able to detect speakers who prefer these forms, we basically managed to identify a group of innovative speakers without making the mistake of drawing misleading conclusions regarding innovative linguistic behavior purely on the basis of demographic data. In other words, we propose that a specific aspect of implicit innovative language attitude may be captured by measuring the automatic (implicit) evaluative reactions (i.e., the implicit attitudes) towards particular innovative linguistic forms, and this evaluative reaction then may be used as an independent variable in the investigation of language change processes.

Naturally, this interpretation of innovative language attitude, that is, the attitude towards a linguistic innovation cannot be understood as a direct predictor of innovative behavior in the production of this particular linguistic innovation, as implicit attitudes are never taken as predictors of behavior without further experimental evidence gathered in psychology either. To understand clearly how the implicit innovative linguistic attitude towards a specific linguistic innovation interacts with the production of the given linguistic form, a considerable amount of further empirical evidence would be needed (which would probably still not clarify the interaction fully, as is experienced in psychology). However, this interaction is actually of no great importance, if the measure of this implicit attitude is interpreted merely as an indicator of tolerance or acceptance of linguistic innovations as such. This way we may operationalize a specific aspect of implicit innovative linguistic attitude that may afterwards be also considered as a possible predictor of (perceptive or productive) innovative linguistic behavior in the case of other independent language change processes as an independent variable. In this sense, a new (and even numerical!) measure reflecting a particular aspect of implicit innovative linguistic attitude may inevitably be an unprecedented advance towards a research method that enables researchers to identify speakers of potentially innovative linguistic behavior. In summary, we argue that through interpreting implicit innovative language attitude as an implicit attitude towards linguistic innovations, the investigation of ongoing language change processes gains a new and useful aspect. Accordingly, we also propose to make an attempt to operationalize this implicit attitude to turn it into a numerically expressed independent variable.

There is a new line of research emphasizing the above described interpretation of language attitude and its role in language change processes. These studies made use of identifying innovative speakers on the basis of the detection of their implicit language attitude towards specific substandard linguistic forms. Mády (2012) aimed at investigating the neutralization of vowel quantity distinction in the case of the Hungarian /o u/ and designated two groups of listeners who were separable based on a test designed to detect language attitude based on instant assessment of language forms. To assess implicit (or rather "semi-conscious") attitude, Mády created a "quasi-offline" explicit evaluation task (here "quasi-offline" means that no RT data was measured, but the response time for the offline answers was limited by the time course of the sound file used in the test): she

concatenated a set of substandard linguistic forms[3] into one audio file, and asked the listeners to spot the "incorrect" forms in it. Supposing that tolerance against substandard linguistic forms, that is, the lack of susceptibility to spot the "errors" is the manifestation of implicit innovative language attitude, Mády used the number of the spotted substandard forms as a measure of implicit conservative language attitude, while innovative attitude was basically defined as the lack of the susceptibility to spot these "errors". The most important result found in the study was that the two groups of listeners separated on the basis of their performance in the implicit task also showed a tendency to perceive the quantity of vowels differently based on their tenseness (i.e., quality difference or spectral difference) which she considered to be a reflection of an ongoing sound change process. Based on these results, Mády also claimed that implicit attitude should be considered as a factor influencing perception, thus it should also be considered as a factor possibly affecting sound change processes. In a second study Mády and Rácz (2013) investigated the sustained /a i u/ (and also the production of /o u/ in embedded words) and found that both young and innovative speakers tended to increase their formant frequencies F1 and F2 for sustained vowels compared to old and conservative speakers (again, innovative attitude was defined on the basis of the evaluative test used in Mády (2012). Therefore, Mády and Rácz claimed that the lowering of the first two formants, i.e., the slightly centralized production of vowels is not purely a result of aging but also of implicit language attitude.

2.5   Aims of the study

The long-term purpose of the authors of the present paper is to follow up on the proposal of Mády (2012) and to develop a test that enables the detection of implicit innovative language attitude directly and reliably. It is suggested that such an implicit attitude test might allow future research addressing the issue of language change to identify innovative speakers independently of demographic data and also of the linguistic variants at question, thus opening up new possibilities in the investigation of language change processes. It is proposed that the reliability of the implicit attitude test used by Mády (2012) can be improved in two aspects. First, the efficacy of bypassing conscious control can be increased by using a completely implicit task (including no explicit evaluation). Second, the efficacy of the detection of innovators may be improved, if we do not detect the respondents' susceptibility to identify speech errors, stable variations or substandard forms (as in Mády 2012), but we include an innovative linguistic form in the implicit method, a form that is assumed to be part of an ongoing language change. Therefore, to develop the implicit innovative language attitude test we propose the following.

First, we suggest the use of the auditory IAT design to detect implicit innovative language attitude directly. The auditory IAT is considered optimal (at least as a first attempt) for two reasons. On the one hand, this paradigm allows for the use of auditory stimuli which is probably the best solution if innovative linguistic forms are to be presented as target items, as innovative variants appear first in spontaneous speech and

---

[3]   i. Variables that show stable variation, such as e.g., the illative case marking suffix *-ba/-be* used in inessive case instead of the inessive case marking suffix *-ban/-ben*; ii. errors like *öbölt* instead of *öblöt* 'bay+ACC'); iii. some non-frequent loan words not used in standard texts, like *sréhen* 'diagonal'; and iv. forms that are characteristic of spontaneous speech, such as the use of the discourse marker *így* approx. 'like'.

may thus appear more natural in an experimental situation if presented to listeners also in speech. On the other hand, we suppose that in the case of the tested opposing target categories, i.e., the innovative and conservative linguistic forms, the conservative is likely to be preferred over the other by conservative users, which is also the rationale behind the IAT paradigm.

Second, we claim that the IAT capturing the implicit attitude of language users towards Linguistic Innovations (referred to as the LI-IAT) should include a linguistic form that is assumed to be part of an ongoing change as test variable. Although, as argued above, this LI-IAT measure may not reflect or predict the production of this particular linguistic innovation we use in the LI-IAT, it should not be considered as a problem, since we only aim to operationalize potentially innovative attitudes towards linguistic innovations.

In line with the second claim, the first practical aim of the study is to carry out an experiment to identify a linguistic variable that is subject to an ongoing change. Based on the test results we create the LI-IAT: we use the innovative ("newer") and conservative ("older") forms of the chosen linguistic variable as contrasting target categories. As attribute categories that are eligible to express valence and thus to evaluate automatic preferences, we decided to opt for the most common good versus bad concept pair.

As a second practical aim of the present study we plan to test the functioning of the newly created LI-IAT, and gather data with randomly sampled informants. The most common numeric output of an IAT is the so called D measure or D score (see Section 2.3) which is generally interpreted as preference of one target category (negative values) or the other (positive values) or no preference (around the value of zero) at a particular target-attribute alignment. However, in the case of the implicit innovative linguistic attitude it is not obvious how to interpret these values. While speakers with conservative implicit attitude may (perhaps straightforwardly) be conceptualized with a preference for the conservative linguistic form, speakers with an innovative implicit attitude may be expected to have a preference for the innovative form, or no preference whatsoever (for further discussion of this problem see the introduction of Section 4). Therefore, we argue that the clarification of the optimal interpretation of the D measure is a long process; in order to understand the nature of the implicit attitude captured by the LI-IAT we need to carry out several thorough empirical studies on language change processes already described for contemporary Hungarian using this new method. In this process we basically plan to map the specific aspect of implicit innovative linguistic attitude that is captured by the LI-IAT to several types of language change processes and also to innovative behavior regarding speech production and perception, in order to assess which type of behavior this specific aspect of implicit attitude (i.e., the D measure of the LI-IAT) can be a reliable indicator of. Consequently, the optimal interpretation of the D scores obtainable by LI-IAT cannot be discussed here. However, we can raise questions regarding possible interpretations of the data and gain basic impressions on the basis of the first data set. Also, we plan to evaluate any design-related biases of the LI-IAT by means of testing normal distribution of the data which is an expected criterion if larger amount of data is collected.

In the following sections we report on the pre-test designed to designate a linguistic variable sufficient for the purposes of an auditory LI-IAT and we also report on the first data that were gathered by means of this newly created LI-IAT in order to explore some of the basic features of the specific aspect of the implicit attitude the LI-IAT's output defines.

3   Experiment 1

To build the LI-IAT, the first step is to identify language variables that are most probably subject to an ongoing language change, and thus have a new or innovative form and an older, conservative one which (in at least some speakers' cases) are competing in language use. To recognize language change in progress, it is assumed that one has to show that a variant has started diffusing in a speech community. For this purpose in the present study we conducted a survey applying the *apparent time construct*, that is, we compared a younger and an older group of speakers' evaluative responses to the linguistic variables at hand, so that we could detect if the evaluative responses differ across age groups.

Since the groundbreaking work of Labov, sociolinguistic literature follows the assumption that the difference in use or in the evaluation of language variables accompanying age is (or may be) an indicator of an ongoing language change (see e.g., Labov 1963, 1966 [2006]). However, it is certainly an issue whether the differences found between age groups should be regarded as "the linguistic change at the community level", or rather "the linguistic change at the individual level", or in other words, whether the change that is bound to the individual's lifespan, i.e., *age grading* (Wagner 2012, 371). The main cause of this differentiation problem is that age grading is indistinguishable from generational change in progress when only apparent time data are available (Wagner 2012).

As it is an old debate in the literature, there is plenty of evidence for and against the reliability of the apparent time construct in the detection of generational change. Nevertheless, drawing on the research of Sankoff (2006), in the present paper we accept the reliability of this method. Reviewing thirteen replication studies of previous research (that used apparent time), Sankoff (2006) found that language change detected in apparent time was confirmed by real time (longitudinal) analysis in all cases. Therefore, she concluded that the apparent time construct is a valid and reliable means of language change detection in most cases. On this basis, although we subscribe to the idea to the fact that apparent time may reflect both age grading and language change to some extent, we assume that a difference found in the language use of older and younger age groups, more specifically, the difference found in the evaluation of particular linguistic variants between older and younger age groups, may be a reliable indicator of an ongoing language change.

As introduced in the previous section, implicit or habitual attitudes are in many cases non-accessible through explicit questions, i.e., questionnaires. Questionnaires, however, are a good means of identifying standards, as they reflect explicit attitude, and they also reflect the social desirability bias: they reflect recently constructed or deliberately formed dispositions and beliefs, and beliefs about what is "socially desirable". Therefore, in Experiment 1 we used a questionnaire to identify the "standards" or explicit bias of speakers in the case of eleven linguistic elements that have two co-existing variants. (For further details, see Section 3.1). The aim of Experiment 1 was to test which of the studied elements proves to be part of an ongoing language change, and thus qualifies as an appropriate variable for the IAT.

3.1   The linguistic variables tested in Experiment 1

In order to find variables that could be tested with the LI-IAT, we assembled a list of 11 structures that were assumed to exhibit variation in MSH either on the basis of our own

observations or according to the literature (most notably Nádasdy 2008). Then we carried out two pre-tests with these variables. On the one hand, we tested the occurrences of their variants in the Hungarian National Corpus (http://clara.nytud.hu/mnsz2-dev/bonito/run.cgi/first_form; for a description, see Oravecz et al. 2014) in order to see whether their distribution shows characteristic patterns. On the other hand, these variables were also tested with the help of informants. In what follows, we will give a brief structural description of each tested variable and summarize the results of the grammaticality judgement test conducted with informants. Throughout the discussion the variants of each variable will be uniformly labeled as "more frequent"/"less frequent", and these labels were assigned on the basis of the corpus study. However, the detailed description of the corpus inquiry and a more thorough structural analysis of the variables would greatly exceed the limits of this paper; therefore, we have made these additional materials accessible in a separate file.[4]

The test asking for native speaker judgement contained the following 11 structures. Each of these are illustrated below with a typical example, with the more frequent variant highlighted in boldface.[5]

(1) Article drop in sentence-initial position:

    a. ***Az ajtók*** *a bal oldalon nyílnak.*
        The door.PL the left side.SUP open.PRS.3PL
    b. *Ajtók a bal oldalon nyílnak.*
        door.PL the left side.SUP open.PRS.3PL
        'Doors open on the left'

(2) Article drop before names of institutions:

    a. *Bemegyek* **a Nyugatiba** *jegyet venni.*
        stop.by.PRS.1SG the Western.ILL ticket.ACC buy.INF
    b. *Bemegyek Nyugatiba jegyet venni.*
        stop.by.PRS.1SG Western.ILL ticket.ACC buy.INF
        'I stop by Nyugati [=Western Railway Station] to buy tickets'

(3) Article drop in non-sentence-initial position:

    a. *Túrázni volt* **a családjával.**
        hiking.INF was the family.3SG.INS
    b. *Túrázni volt családjával.*
        hiking.INF was family.3SG.INS
        'She went hiking with her family '

---

    [4] http://full.btk.ppke.hu/index.php/FULL/article/view/50/61
    [5] List of glosses used in the present paper: ACC = accusative; DAT = dative; FUT = future, ILL = illative; INF = infinitive; INS = instrumental; MOD = marker of epistemic modality; PL = plural; PRS= present; PST = past; PV = preverb, SUP= superessive.

(4) Presence or absence of the subordinator *hogy* in sentences with the verb *lehet*:

    a. **Lehet,**   **hogy**   *ő*   *is*   *eljön.*
        be.MOD.PRS.3SG   that   (s)he too   come.PRS.3SG
    b.   *Lehet,*   *ő*   *is*   *eljön.*
        be.MOD.PRS.3SG (s)he too   come.PRS.3SG
        'It's possible that (s)he comes over, too'

(5) Presence or absence of the subordinator *hogy* in sentences with the sentence adverbial *természetesen* 'naturally':

    a. **Természetesen** *pontosan*   *érkezik.*
        Naturally   on.time   arrive.PRS.3SG
    b.   *Természetesen,*   *hogy* *pontosan*   *érkezik.*
        Naturally,   that on.time   arrive.PRS.3SG
        'Naturally, (s)he arrives on time.'

(6) Presence or absence of the subordinator *hogy* in sentences with the sentence adverbial *valószínűleg* 'probably':

    a. **Valószínűleg** *holnap*   *János*   *is*   *ott*   *lesz.*
        probably   tomorrow John   too   there be.FUT.3SG
    b.   *Valószínűleg,* *hogy* *holnap*   *János*   *is*   *ott*   *lesz.*
        probably   that tomorrow John   too   there be.FUT.3SG
        'Probably John will be there tomorrow'

(7) Presence or absence of the subordinator *hogy* in sentences with the sentence adverbial *nyilván* 'obviously':

    a. **Nyilván**   *az*   *én*   *ebédemet*   *ette*   *meg.*
        obviously   the   my   lunch.1SG.ACC   eat.PST.3SG<3SG PV
    b.   *Nyilván,*   *hogy* *az*   *én*   *ebédemet*   *ette*   *meg.*
        obviously   that the   my   lunch.1SG.ACC   eat.PST.3SG<3SG PV
        'Obviously it was my lunch (s)he ate'

(8) Optional *hogy*-deletion:

    a. *Azt*   **hittük,**   **hogy** *itt* *tilos*   *a*   *dohányzás.*
        that.ACC   believe.PST.3SG<1PL that here forbidden the   smoking.
    b.   *Azt*   *hittük,*   *itt* *tilos*   *a*   *dohányzás.*
        that.ACC believe.PST.3SG<1PL   here forbidden the   smoking
        'We thought that smoking was forbidden here.'

(9) Unmarked object:

    a. *Levittem*   *a*   **kutyám-at**   *sétálni.*
        take.PST.3SG<1SG the   dog.1SG-**ACC**   walk.INF
    b.   *Levittem*   *a*   *kutyám*   *sétálni.*
        take.PST.3SG<1SG the   dog.1SG   walk.INF
        'I took my dog for a walk'

(10)   Compounds of the type noun+verb:

   a.   *Holnap*   *egész*   *nap*   **ügyeket**   **intézek**.
        tomorrow   whole   day   affair.PL.ACC   manage.PRS.1SG

   b.   *Holnap*   *egész*   *nap*   *ügyintézek.*
       tomorrow   whole   day   affair.manage.PRS.1SG
       'I will be managing affairs tomorrow the whole day.'

(11)   Unmarked plurality of the possessor on the possessee:

   a.   *A*   *gyerekeknek*   *nincs*   **étvágy-uk.**
        the   children.DAT   not.exist   appetite-**3PL**

   b.   *A*   *gyerekeknek*   *nincs*   *étvágy-a.*
       the   children.DAT   not.exist   appetite-3SG
       'The kids have no appetite'

The first three phenomena represent different types of article drop. The Hungarian definite article *a/az* emerged during the Old Hungarian period through the functional split of the distal demonstrative pronoun *az* 'that', a grammaticalization change that was attested in various languages (Heine & Kuteva 2004, Harris & Campbell 1995, Givón 2001). Its use displays gradual extension during Old Hungarian, appearing in more and more grammatical contexts where it marks definiteness (I. Gallasy 1991, 1992, Egedi 2014).

However, in the test we looked at examples which could (theoretically) be interpreted as a change in the opposite direction:[6] the article does not appear in environments in which it would be compulsory, or at least preferable in MSH. One such example was observed by Nádasdy (2008), see (12).

(12)   *Nyugatiba*   *bejössz* […]
      Western.ILL   drop.in.PRS.2SG
      'You come into the Western (Railway Station)'

With respect to example (12), Nádasdy claims that while this form does not confirm to MSH (and one reason for this is the lack of the article before *Nyugati)*, it is still generally observable in spoken language, and the environment that might trigger article drop is the sentence-initial position of the noun. We decided to test two aspects of this phenomenon, one of them being sentence-initial position (following Nádasdy), and the other the category of names of institutions (whether these show some tendency to be used without an article). In addition, we added sentences including nouns in medial position with or without articles as a third type. Naturally, the latter is too broad a category to be tested in general, therefore, we selected sentences to be tested in which we happened to observe article drop ourselves (e.g., the article-less version of example (3) was observed by the authors in spontaneous communication, although as native speakers we would strongly prefer the variant with the article).

---

   6   However, it is almost impossible to investigate whether in these specific instances the use of the article had already been general, and the drop of the article is a further change, or there has always been a significant amount of variation, and, due to so far unknown reasons, the more archaic, article-less pattern started to spread in MSH.

Sentences 4-8 all target the presence or absence of *hogy* 'that'. This is the most general subordinator in Hungarian, a complementizer that developed presumably during the Proto-Hungarian period through the reanalysis of the question word *hogy* 'how' in embedded questions or, alternatively, of the pronominal adverbial *hogy* 'as, the way that', which were homophonous at the time of this change (in MSH, the latter is *ahogy;* for a general description of the grammaticalization process, see Haader 1991). Investigation of the earliest sources reveals that by Old Hungarian, the complementizer appears in all of the functions that it has in MSH (Haader 1995, Bácskai-Atkári & Dékány 2014). However, it was not an obligatory marker of finite subordination then, and its use is not obligatory in MSH, either; for an overview of the conditions of *hogy*-deletion, see Kenesei (1992, 673-679).

Out of the three tested environments, one contains examples where the complementizer could be freely omitted in MSH, as in example (8) above. Another tested type was the deletion of the complementizer in sentences in which the verb of the main clause was *lehet* ('it may be'). In contrast to the previous type, the subordinate clause of this matrix verb would necessarily be headed by an overt complementizer (i.e., *hogy*) in MSH. Still, especially in non-formal registers, *hogy*-deletion seems to be spreading. Finally, the last set of sentences that focus on the presence or absence of *hogy* (phenomena 5, 6, and 7 in the test) do not contain matrix verbs at all. The striking feature of the given structure is that there is an adverbial that seems to govern a subordinate clause in this case. This structure was first described in the seventies (see e.g., E. Abaffy 1976), but some highly sporadic instances can be attested already in Late Old Hungarian (Haader 2001). Owing to the truly unique nature of this pattern, there were quite a few attempts to provide a structural analysis and/or an account of its development; these are outlined in the supplementary material, and a thorough review is available in É. Kiss (2010).

Phenomenon 9 pertains to object marking. In MSH, all objects are marked with the accusative suffix *-t* (13a-b).

(13) a. *jön*   *a*   *hajó*
          come.PRS.3SG   the   ship
          'the ship is coming'
      b. *látom*   *a*   *hajó-t*
          see.PRS.3SG<1SG   the   ship-ACC
          'I see the ship'

There is one regular exception: its use is not obligatory with nouns marked with a Sg1 or Sg2 possessive suffix (14b, cf. 14a).

(14) a. *jön*   *a*   *fia-m*
          come.PRS.3SG   the   son-1SG
          'my son is coming'
      b. *látom*   *a*   *fia-m*
          see.PRS.3SG<1SG   the   son-1SG
          'I see my son'

As for the potential factors motivating the speakers to use the unmarked or the marked alternative, their choice seems to be grammatically unconstrained, and neither of the forms is stigmatized or bound to registers. Diachronically, the unmarked accusative

in this case is an archaic feature with parallels from e.g., Eastern Mansi (see Virtanen 2013).

As opposed to the age-old variation observable in phenomenon 9, phenomenon 10 seems to be a recent innovation: a type of compounding in which the first stem of the compound is a noun, and the second is a verb, e.g., *apróhirdet* 'to post small ads', *bájcseveg* 'to do small talk', *agymos* 'to brainwash' (Kiefer 2000, 531). At first sight, these seem to consist of a verb and one of its arguments, the latter losing its appropriate morphological marker owing to the process of compounding, e.g., *város-t$_{ACC}$ néz* 'to go sightseeing', lit. 'watches the city'→ *városnéz*. However, this analysis would not be correct: these noun+verb type of compounds are backformations from pre-existing derivations in which the last stem is an action noun formed with the suffix *-ás/-és*, e.g., *apróhirdetés, bájcsevegés, agymosás*. Kiefer (2000, 531) points out that sporadically, some compounds of this type may arise straight from a verbal phrase due to analogy, but this pattern of compounding is not (yet) productive in Hungarian. As these forms are morphologically transparent, all native speakers can interpret (and, theoretically, can produce) the innovative, noun+verb compounds.

Finally, phenomenon 11 is also an instance of long co-existing variants. In MSH possessive constructions the possessor (in all person-number variations) can be either in the nominative (15a) or in the dative (15b). As opposed to the optionality of markedness described above in the case of object marking, variation here is not arbitrary: the different cases are due to the different structures of these two patterns (for a detailed structural description of the possessive constructions, see Szabolcsi & Laczkó 1992).

(15) a. *Péter   ház-a*
         Peter   house-3SG
         'Peter's house'
      b. *Péter-nek a   ház-a*
         Peter-DAT   the   house-3SG
         'Peter's house'

If the plural third person possessor is encoded as a noun in the dative as in (16), the possessee can either have a singular third person possessor marker as in (16a) (this would be the general type with nominal possessors in the nominative), or it can be marked with a plural third person possessor marker as in (16b) (which would be the general type with pronominal possessors).

(16) a. *a   fiú-k-nak   a   ház-a*
         the   boy-PL-DAT   the   house-3SG
         'the boys' house'
      b. *a   fiú-k-nak   a   ház-uk*
         the   boy-PL-DAT   the   house-3PL
         'the boys' house'

The variation that is characteristic of MSH (as in example (16)) is already present in Old Hungarian (Korompay 1991, 1992), and, naturally, these two forms must have also co-existed between these two periods (Old and Modern Hungarian) as well. This is shown by the fact that, according to the survey of the handbook of normative linguistics (Grétsy & Kovalovszky 1980, 350), representatives of the language reform movement (taking place between 1790-1820 approximately) propagated the exclusive use of the

agreeing form both with the nominative and with the dative possessor, meaning that they must have been aware of the presence of competing variants, i.e., the agreeing and the non-agreeing form. It is also instructive to see how the authors of the handbook of normative linguistics interpret this variation concerning MSH: they note that the use of the non-agreeing form is spreading in MSH, but they also recommend the use of the plural-marked possessee with dative-marked possessors.

3.2   Participants, procedure

The 11 linguistic phenomena described in the previous section were represented by 5 examples in the grammaticality judgment test resulting in 55 test sentences which were presented in pairs and in a randomized order with 55 distractor sentences (that consisted of dialectal forms, cases of free variation in vowel harmony and ungrammatical sentences). 54 adult participants (12 male, 42 female) clustered in three age groups provided data in the experiment: *(17 ≤) x ≤ 21 year olds* (14 participants), *21 < x < 50 year olds* (29 participants), and *50 ≤ year olds* (11 participants). The groups were targeted directly through online social networks and mailing lists.

It was assumed that the evaluation of language variables, i.e., the recognition of the language variables as "correct" or "incorrect" forms should not necessarily require direct introspection (or any reflection on one's own language use). It was suggested that evaluation of correctness automatically involves one's own language norms in naïve speakers without further suggestion. Additionally, it was also desirable to reveal the relation of the two competing variants very clearly, that is, to see whether the acceptability of one variant arises at the expense of the other, or whether the two variants are accepted simultaneously and to categorize the fine scales of acceptability (obtainable for instance by means of a Likert-scale) into clear groups of acceptable and non-acceptable language forms, as this is an inevitable requirement for the use of the IAT. Last but not least, it was important to make sure that participants do not assess any other features of the tested sentences, but the phenomena in question. These aims were achieved by the following features of the questionnaire used in the experiment.

First, the attention of the participants was focused on the tested linguistic phenomena by presenting the contrasting sentences in pairs. Second, we excluded any aspect that might elicit any self-report (on the participants' own language usage) and its conflict with the participants' supposed language norms, while forcing them to give categorical answers: the task was narrowed down to answering the simple question which of the two sentences the participants found to be correct (possible answers: a) *first*, b) *second*, c) *both – with the same meaning*, d) *both – with different meanings,[7]* or e) *I do not know*). At the end of the questionnaire the participants answered some demographic questions (age, gender, education, etc.) and they could also add further comments. In the analysis no other demographic features but age were directly used, the rest of the responses were merely checked to assess the reliability of the grammaticality judgments. The survey was administered through the internet.

---

[7]   The categories *both – with the same meaning* and *both – with different meanings* were not considered to be reliably reported on by the participants, thus these were collapsed into one category later on in the analysis.

3.3   Results

Figure 2 summarizes the results of the sentence evaluation task in the two outermost age groups (≤ *21 year olds* and *50 ≤ year olds*).
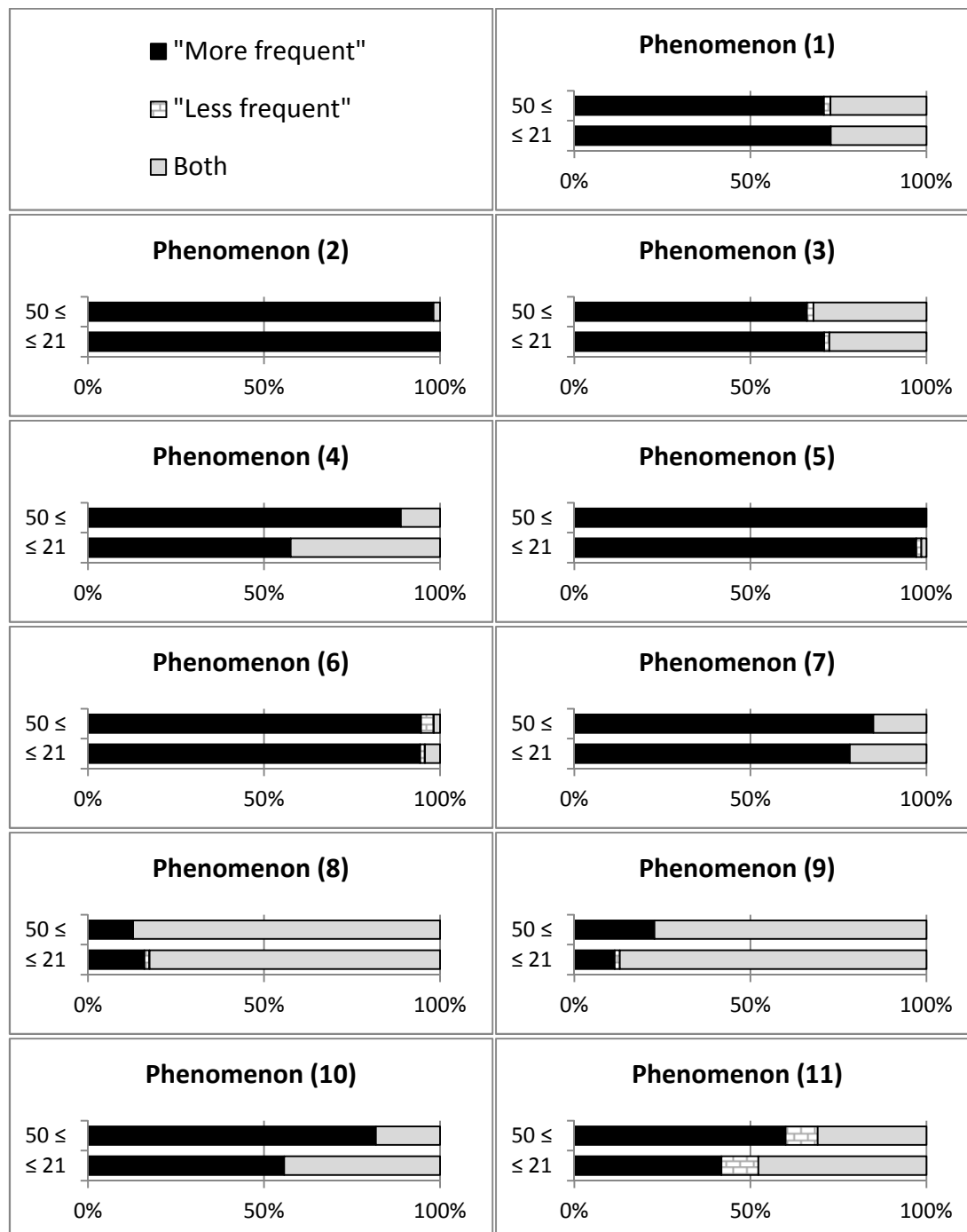

Figure 2: *The evaluation of the 11 phenomena (i.e., 22 contrasting sentence pairs) in the youngest and oldest age group (answering the question "Which of the two sentences do you find to be correct?"). The label "more frequent" refers to the results of the corpus analysis*
(see Section 3.1 and the Supplementary Material)

Concerning the potential spread of a less frequent form, clear evidence for this was found only in three cases: (4) *("Lehet, hogy… /Lehet…")*, (10) (e.g. *"ügyeket intézek/ügyintézek"*), and (11) (e.g. *"a gyerekeknek nincs étvágyuk/étvágya"*). At first it appears that (11) is very similar to (4) and (10). However, its case is most probably different from the other two phenomena, since (11) is the only phenomenon in which the exclusive acceptance of the less frequent form can be detected in a notable percentage (in almost 10% in both age groups).

Although it is not visible in the graphs, according to post-questionnaire comments of some of the participants (10) also turned out to be an "outlier" in some sense. It is clearly seen in Figure 2 that the younger group accepts the "less frequent" variant of (10) to a larger extent than the older group. However, according to their remarks, the participants do not accept the "less frequent" variant due to it being "innovative", but rather because of a "difference in style" (relative to the phrasal expression). Namely, some participants noted that they found the "less frequent" forms of (10) to be funny expressions which they regard as correct if used only in jest (in joking, teasing or other highly informal or intimate interactions). The perceived stylistic difference was noted by participants from both age groups.

Phenomena related to article deletion (1), (2), and (3) seem to behave similarly to each other with respect to grammaticality judgments. However, (2) shows a different pattern than that of (1) and (3), i.e. proper noun phrases behave differently from common noun phrases with respect to the deletion of the definite article. It appears that deletion of the definite article of proper noun phrases (where the noun refers to institutions/buildings) is the least accepted phenomenon observable in everyday speech among the tested variables, along with (5), (6) and (7), i.e. the cases concerning the presence or absence of the subordinator *hogy* in sentences with a sentence adverbial.

The results for (8) and (9) suggest that optional *hogy*-deletion and marked versus unmarked object are examples of variables with widely accepted variants (the less frequent forms and the more frequent forms were equally accepted by 80% of the informants in both age groups), therefore, these forms most probably display stable variation.

As phenomena (10) and (11), as discussed above, turned out to be more challenging to interpret, it was only phenomenon (4) that was chosen for further analysis. The data were re-grouped into 3 age groups ≤ *21 year olds, 22 < x < 45 year olds,* and *45 ≤ year olds* that are more well-balanced in number (14, 24, and 16 participants respectively), to control for the effect that may be introduced by count differences of the original groups. As can be seen on Figure 3, the less frequent and most probably also innovative form *("Lehet")* of phenomenon (4) is regarded as correct in less than 10% in the oldest age group *45 ≤,* while both variants are accepted in more than 55% in the group *21 < x < 45*, as well as in the youngest age group *(≤ 21).*
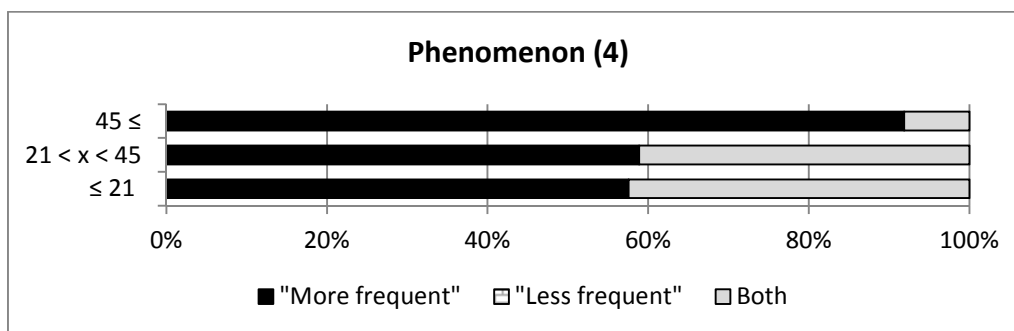
Figure 3: *The evaluation of the sentences of Phenomenon (4) (answering the question "Which of the two sentences you find to be correct?") if the data are grouped into 3 age groups 21 ≤ year olds, 22 < x < 45 year olds, and 45 ≤ year olds that are more well-balanced in number (14, 24, and 16 participants respectively)*

3.4 Conclusions

We may draw the conclusion that, as expected, only some of the tested phenomena showed a shift in the apparent time, i.e., only some of the phenomena proved to be a diffusing innovation spreading. Concerning phenomena (1), (2), and (3), neither the corpus query nor the test data seem to confirm the authors' observations that article drop would be an innovation that is spreading in the speech community. However, we would still like to suggest that article drop in certain structural configurations is an innovative form, the presence of which is attestable in the vernacular. It should be kept in mind that while the grammaticality test was conducted in written form, these phenomena and linguistic innovations generally and mostly are observable (first) in informal and spontaneous speech. Therefore, it is very probable that our findings, i.e., the fact that informants do not seem to accept this "innovative" variant in written form, are only indicative of the acceptability of this variant in the given modality, i.e., writing. On the very same basis, however, "less frequent" variants that show greater acceptance in the younger age group under these circumstances may be more reliably considered as linguistic innovations, since, again, they provide evidence for relatively high percentages of acceptance despite the unusual modality.

Based on the reasons discussed above, the second, and most important conclusion of Experiment 1 is, that the *'Lehet…'* variant of phenomenon (4) *("Lehet, hogy…/Lehet…")* is a linguistic innovation that clearly shows greater acceptance in the younger age groups.

According to an evolutionary account (see e.g. Labov 1972, Croft 2000) the differences found between (4) and (11), or (10) and (11) may be interpreted as follows. Consider phenomenon (11) first. The data show that the two competing variants are not just equally accepted, but for some speakers, the new variant has already overridden or replaced the other. If the "less frequent" variant is taken to be an innovation, this reflects that language change has progressed further with regard to (11) than with other phenomena. If, however, one considers this issue from a diachronic point of view, the putative innovative (less frequent) variant is in fact the more archaic member of this pair. In view of the ubiquitous but so far mostly unexplored fluctuation in the distribution of the two forms in throughout the history of Hungarian, we found it necessary to avoid categorization in this case.

The judgments of (8) and (9) provide examples of stable variation. Similar rates of acceptability of "less frequent" and "more frequent" variants support this interpretation.

Phenomenon (10) also showed differences in the acceptance in the apparent time. However, it was also noted that the "less frequent", or, so to speak, innovative variant of this phenomenon was also perceived very different in style or register.

On the basis of the findings discussed above, we concluded that phenomenon (4) may be regarded as the optimal input construction for the LI-IAT, as only this variable showed a clearly identifiable innovative and conservative variant. Therefore, throughout the rest of this paper, as well as in the construction of the LI-IAT, we focused on phenomenon (4), *"Lehet, hogy…/Lehet…"*. In the next section we describe how the auditory LI-IAT is built and present some data gathered by this new tool.

4   Experiment 2

Traditionally, IAT is (mainly) used for the detection of stereotypes and prejudice against such stereotypical groups as African Americans/white Americans, overweight people/thin people, or heterosexuals/homosexuals. Therefore, the traditional IAT can make use of the stereotypically available representations of the opposing groups: faces, names or words that are (again, stereotypically) associated with the contrasting concepts. In a language attitude IAT, however, where innovative and conservative language forms (and language users) are to be assessed, there is no obvious conceptualization or visualization available that would unequivocally identify the contrasting variables. There is a possibility to use arbitrary visualizations, like pictures of younger "innovative" speakers and older "conservative" speakers. However, it can easily introduce some undesired artifact or bias to the results that cannot be controlled for in any way, since it is most probably the positive/negative attitude *towards the particular faces* that would be measurable in a design where compatibility (i.e., the association of an innovative form with the assumed innovative speaker's picture) cannot be taken for granted. One side of the problem is also known as the 'halo effect', i.e., the problem that beauty biases our perception in a manner that we make choices or decisions in favor of the thing we are attracted to. Another side of the problem using faces for visualization of speaker groups is complexity, as people are a collection of features and also of several stereotypes (in connection with ethnicity, gender, hair style and color, age, skin, etc.) at the same time. As a result, the stereotype activated in a given situation (and for a particular listener) may not be easily inferred and cannot be easily controlled for either (see MacFarlane & Stuart-Smith's review on the topic in MacFarlane & Stuart-Smith 2012). Pantos and Campbell-Kibler provided positive evidence for the possibility to use metalinguistic expressions to mark contrasting variants, like labels of *American* versus *Foreign* (Pantos 2010) or –IN versus –ING (to express differences between examples, like *talkin'* vs. *talking*) (Campbell-Kibler 2012). Luckily, the use of abstract metalinguistic expressions is also a good solution to substitute arbitrary visualizations in the case of the present study.

In order to identify the proper meta-linguistic category labels for the contrasting linguistic forms of the LI-IAT, we conducted a pre-test involving 46 participants (not reported here in detail). In this first attempt the category labels *newer* and *older* were tested. The result showed that these category labels do not sufficiently cover the innovative and conservative variables, as participants assigned these labels to both of the members of the contrasting pairs equally (in approx. 50%). Therefore, for the construction of the LI-IAT we finally opted for using a more direct representation of the variables at hand by

assigning the conservative and the innovative target label categories LEHET, HOGY (POSSIBLE THAT) and LEHET (POSSIBLE) to the members of the contrasting pair in the LI-IAT.

The aim of Experiment 2 was to create the LI-IAT with the variable *"Lehet, hogy…/Lehet …"* selected in the previous experiment, to observe the distributional characteristics of its outcome measure, and investigate the characteristics of the captured aspect of implicit attitude. We intended to evaluate any design-related biases of the LI-IAT that can be observed, and to raise questions regarding possible interpretations of the data (D measures) in the measurement of innovative implicit language attitude. Traditionally, D measures are interpreted as preference for one category over the other ($D \geq .15$ and $D \leq -.15$) and the lack of preference ($-.15 < D < .15$) which will also be analyzed here, in Experiment 2. However, in the present paper we must leave it an open question to designate where the exact boundary between the two categories of speakers (conservatives and innovators) should be drawn on the D measure continua (from -2 to $+2$). It is clear that this question should not be (moreover, cannot be) answered based on IAT data alone, but needs an extensive amount of follow-up research: one must correlate the LI-IAT data with many sets of linguistic data on variants that may be considered innovative and analyze these correlations very thoroughly. For the time being, we chose the following course of action. First, we merely adapted the traditionally used effect size criteria and designate conservatives by having at least a weak preference for the conservative variant of the language variable ($D \geq .15$), while we designate innovators lacking this preference ($-.15 < D < .15$) or having preference for the innovative variant ($D \leq -.15$). Second, we investigated interrelations of implicit attitude (detected by the LI-IAT) with age, and gender. Third, we studied interrelations between explicit and implicit attitude (detected by the LI-IAT) to see if we find converging implicit and explicit measures as observed by Pantos (2010) and Pantos and Perkins (2013). Fourth, we assessed if there is any bias observable in the functioning of the LI-IAT by testing whether the data show normal distribution.

It is very important to emphasize here again that in the present paper and in this adaptation of the IAT, implicit innovative linguistic attitude is basically defined as implicit (or automatic) perceptual preference, and that in this study we do not intend to create a test that predicts a speaker's use of the linguistic variant built in the LI-IAT. Additionally, we also cannot venture to clarify the exact interpretation of D measures without a substantial amount of further empirical work. Finally, it is important to note that auditory IAT is an already validated and even standardized technique, and thus its adaptation using new variables (that is, new categories and items) should also be regarded as valid.

4.1 Methods

To gain a first impression on the nature of the output of the LI-IAT, we collected and explored data from 40 adult informants (11 men, 28 women, 1 unknown due to missing data), including 12 informants who also participated in Experiment 1. The exploratory analysis involved the followings.

First, it was tested whether the obtained D measures follow normal distribution by the use of Shapiro-Wilk test of normality in R (R Core team, 2013).

Second, the implicit measures (D measures) were compared with two types of explicit measures: *explicit evaluation* and *self-reports*. The first explicit attitude measure was derived from sentence evaluation responses for the five example sentences of the

variable *"Lehet, hogy…/Lehet…"* obtained in Experiment 1. The evaluative responses were converted into (nominal) explicit attitude measures on the basis of the number of accepted variants: if the participants preferred the conservative variant over the innovative variant exclusively more than 2 times (out of the 5 cases), they were labeled as conservative speakers. Otherwise, they were labeled as innovative speakers. We contacted again those volunteers who provided their email addresses for further research in Experiment 1 and asked them to participate in Experiment 2 as well. Eventually, 12 volunteers (4 men, 8 women) participated again, thus their data were analyzed in the first comparison (i.e., comparison between evaluative responses as explicit measures and implicit measures). The second type of explicit measures consisted of self-reports that were obtained in Experiment 2: after finishing the IAT, participants were asked to answer some basic demographic questions again, as well as two questions on their language attitude used previously by Mády (2010). Participants were asked to indicate their agreement with the following statements by means of bipolar questions (possible answers: *yes* or *no*):

1. *"Mindig nagyon figyelek arra, hogy szépen, érthetően beszéljek."*
    'I always pay particular attention to use correct pronunciation'
2. *"Szerintem a túlzottan pontos kiejtés mesterkélt, nem illik a mindennapi beszédhez"*
    'I think that overly accurate pronunciation is mannered and is not appropriate for everyday speech'

Agreement with the first statement was regarded a statement of conservative explicit attitude, while agreement with the second statement was regarded a statement of innovative explicit attitude. Participants who agreed or disagreed with both statements were excluded from the comparison. There were 27 participants in Experiment 2 who provided different answers for the two explicit questions, thus only this group of 27 informants was eligible for this comparison. In the analysis no other demographic features but age and gender were directly used, the rest of the responses were merely used to cross-check the reliability of data collection.

Third, to describe age and gender distribution of conservative and innovative groups, we grouped the data of the 40 participants into conservative ($.15 \leq D$) and innovative ($D < .15$) speakers.[8]

In the present study, when creating the LI-IAT, the traditional IAT design was used, five block types and seven blocks total: in Block 1 participants sorted target items (POSSIBLE THAT vs. POSSIBLE) and learnt the hand-side assignment; in Block 2 they sorted attribute items (GOOD vs. BAD); in Block 3 target and attribute items were mixed, to be sorted simultaneously (POSSIBLE THAT OR GOOD vs. POSSIBLE OR BAD or vice versa, it was balanced equally in the test which of the two possible combinations came first in a particular informant's case); in Block 4 the task of Block 3 was repeated; in Block 5 participants learned the inverse hand-side assignment of target categories; in Block 6 the participants were to sort items according to the second combination of attribute + target categories; and finally in Block 7 the task of Block 6 was repeated (see Figure 1).

---

[8]   It is emphasized here again that in the present study, innovative speakers were designated according to the traditional effect size criteria and on the basis of lacking the implicit perceptual preference for the conservative variant.

For attribute categories the labels JÓ 'good' and ROSSZ 'bad' were used and words with positive and negative valence[9] were selected as items (8 item per attribute category). For attribute and target items used in the LI-IAT see Figure 4.

Items belonging to the attribute categories were presented visually as words appearing in the middle of the screen. The target items consisted of 8 sentence pairs in which the members contrasted only in the linguistic variable *"Lehet, hogy…/Lehet"*. The target stimuli were recorded previously, uttered by a female speaker capable of producing natural sounding, uniform intonation across the two (innovative and conservative) versions of the same sentence. Hence the target items were presented as audio stimuli, while during each auditory trial the visual prompt *"Hang – Kattintson!"* 'sound – please click!' focused the attention of the participants to the recording and also reminded them to sort the item into the correct category. As usual in an IAT, to calculate *D* measures *congruent* trials and *incongruent* trials had to be designated. Since according to the authors' impression conservative language attitude is a commonly experienced explicit attitude bias, conservative attitude (i.e., strong association between the conservative form POSSIBLE THAT and the attribute category GOOD) was designated as congruent trial. To control for any seriality effect, the order of congruent and incongruent blocks was randomized according to the randomly generated participant ID: for even numbers congruent trials came in Block 3 and 4, while for odd numbers incongruent trials came in Block 3 and 4.

The LI-IAT tests were administered via internet; stimuli presentation and recording of the data were both controlled by a script created in Inquisit 4 (Millisecond Software LLC). In the Inquisit 4 software, the measurement precision of RT is ensured by Java Network Launching Protocol (JNLP) which installs a small Java program on each participant's computer in order to run the actual study task. Therefore, RT data are not gathered directly through the internet, but on each participant's computer and they are only transferred to the web store after the task is completed. As a result, this method enables the implementation of reaction time tasks with reasonable precision (Maniaci & Rogge 2014).

D measures were calculated according to Greenwald et al. 2003 (see also Section 2.3 of the present paper); trials greater than 10,000 ms were automatically deleted. Since attribute items were presented visually (as strings), and target items were presented auditorily, modality switch was introduced into the paradigm. Please note, however, that all the previous auditory IAT studies used the same design (Vande Kamp 2002, Pantos 2010, Campbell-Kibler 2012), thus auditory IAT was originally validated with the modality switch condition.

---

[9] Although in the present experiment we adapted the attribute items that were used also in other studies (see e.g., the Project Implicit at https://implicit.harvard.edu/implicit/) and the synonyms of these words, we also conducted an informal pre-test on the valence of these items (i.e., the intrinsic attraction or aversion toward the concepts) via an email survey with 25 participants. In this test the agreement on the valence of the attribute items was basically 100% and the evaluation was also congruent with our classification of these items (with the negligible exception of 2 unexpected responses in the case of *fájdalom* 'pain', and 1 unexpected response in the cases of *gonosz* 'evil', *félelem* 'fear', *gyötrődés* 'torment', *szörnyű* 'horrible').

Categories (and category labels)	Items
Target	
LEHET 'possible'	*Lehet szerzek én is.* 'It is possible (that) I will get one as well.' *Lehet megveszem.* 'It is possible (that) I buy it.' *Lehet kutya volt.* 'It is possible (that) it was a dog.' *Lehet nincs is.* 'It is possible (that) there isn't any.' *Lehet bent maradok.* 'It is possible (that) I stay inside.' *Lehet rosszul láttam.* 'It is possible (that) I had it mistaken.' *Lehet kitöltöm én is.* 'It is possible (that) I will also fill it.' *Lehet így marad.* 'It is possible (that) it stays at is.'
LEHET HOGY 'possible that'	*Lehet, hogy szerzek én is.* *Lehet, hogy megveszem.* *Lehet, hogy kutya volt.* *Lehet, hogy nincs is.* *Lehet, hogy bent maradok.* *Lehet, hogy rosszul láttam.* *Lehet, hogy kitöltöm én is.* *Lehet, hogy így marad.*
Attribute	
JÓ 'good'	*Vidám* 'joyful' *Csodálatos* 'wonderful' *Béke* 'peace' *Szabadság* 'freedom' *Boldog* 'blissful' *Öröm* 'happiness' *Szépség* 'beauty' *Szeretet* 'love'
ROSSZ 'bad'	*Gonosz* 'evil' *Rettegés* 'horror' *Gyötrődés* 'torment' *Szörnyű* 'horrible' *Szomorú* 'sad' *Félelem* 'fear' *Borzasztó* 'terrifying' *Fájdalom* 'pain'

Figure 4: *Target and attribute categories and items used in the LI-IAT*

4.2   Results

According to the Shapiro-Wilk test of normality the *D* measure data of the 40 participants followed normal distribution (see Figure 5). As at large number of data normal distribution is expected in an IAT paradigm, this result suggests that the newly created LI-IAT does not introduce a design-related bias or anomaly to the data.
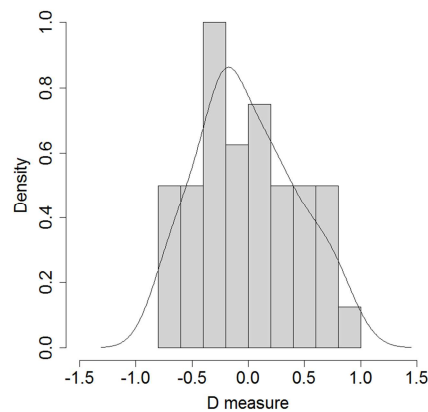
Figure 5: *Distribution of D measures of 40 participants in the LI- IAT*

We obtained explicit attitude measures in both Experiment 1 (sentence evaluation) and Experiment 2 (self-report). We compared both of these data sets with the implicit attitude measurements (*D* measures) obtained in Experiment 2. Figure 6 shows the number of agreement and disagreement between explicit and implicit measures if the separating line between conservative and innovative groups in the implicit test is drawn according to the traditional effect size criteria, i.e., at the weak preference of the conservative form (that is, at .15).
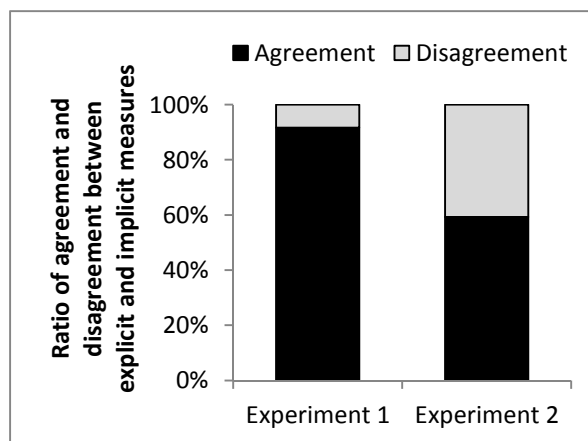

Figure 6: *Interrelations of implicit and explicit attitude measurements as a function of the different types of explicit measures: Experiment 1 – sentence evaluation (12 participants); Experiment 3 – self-report (27 participants)*

There were 12 participants in Experiment 1 from whom it was possible to collect data again, thus the relationship between sentence evaluation and implicit attitudes was assessed based on these 12 informants' data. There were 27 participants in Experiment 2 who provided different answers for the two explicit questions, thus the relationship between self-report and implicit attitudes was assessed based on these 27 informants' data. As opposed to the findings of Pantos (2010) and Pantos and Perkins (2013) (and e.g., Mády 2012 who measured implicit attitude towards substandard linguistic forms), explicit and implicit attitude did not differ to a great extent in our data (i.e, the ratio of agreement is relatively high in both comparisons, see Figure 6). However, differences

between the two explicit measures, i.e., between evaluative responses (Experiment 1) and self-reports (Experiment 2) can also be inferred from the data. In the case of twelve participants providing data in both experiments, explicit attitude measures differed highly between the two experiments (in 58%), whereas agreement was found in only 42% between evaluative responses and self-reports.

The interrelations observable in Figure 7 imply that among all the mismatches between explicit and implicit measures it is *implicit innovative language attitude* that is paired mostly with *explicit conservative attitude* (in Figure 7 see *"Denied innovativity"*), while *implicit conservative attitude* is rarely "denied" (i.e., it is rarely matched with explicit innovative attitude, in Figure 7 *"Denied conservativity"*). Also, mismatches of implicit conservative attitude and explicit innovative attitude can only be observed in the self-report measures condition (data from Experiment 2).
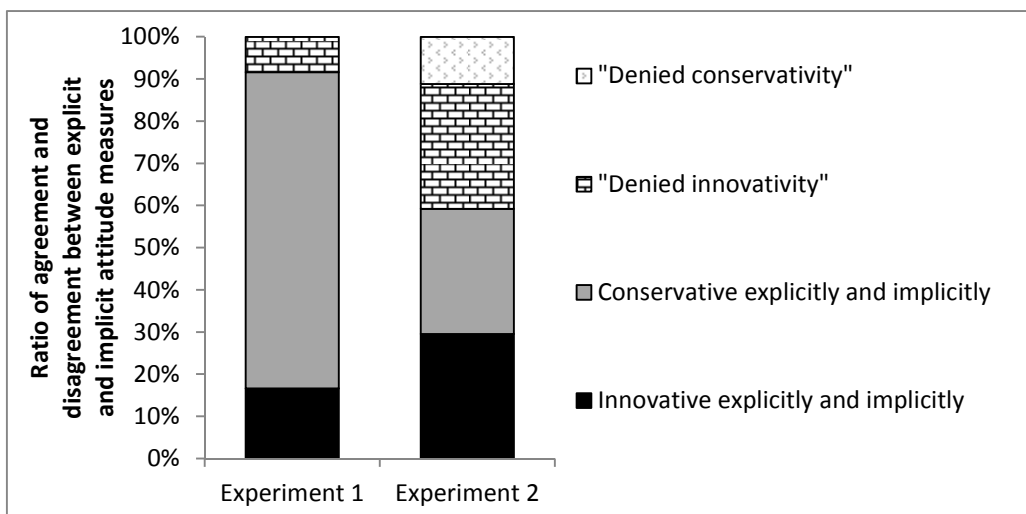

Figure 7: *Interrelations of implicit and explicit attitude measurements as a function of the different types of explicit measures: Experiment 1 – sentence evaluation (12 participants); Experiment 2 – self-report (27 participants)*

Figure 8 shows the age and gender distribution of innovative and conservative speakers designated according to the traditional effect size criteria of the IAT design (separating line of .15, i.e., weak effect in the *conservative* direction). As expected, innovative speakers are more frequent in the younger adult groups, while conservative speakers are found in almost all age groups. As opposed to the expectations, both innovative and conservative groups consist mainly of females, that is, we cannot say that females are more frequent in the innovative group or that females are more likely to belong to the group of innovators. The relations of gender in the two (innovative and conservative) groups also reflect the overall gender hierarchy in the sample: in the randomly sampled subset of the study the male–female ratio is 28% to 72% (similarly to the ratio observable in Figure 8). If we compare the percentage of innovative women (68% of all women) and innovative men (64% of all men), the percentage difference between the two is 4%, meaning that the difference between gender groups is negligible.
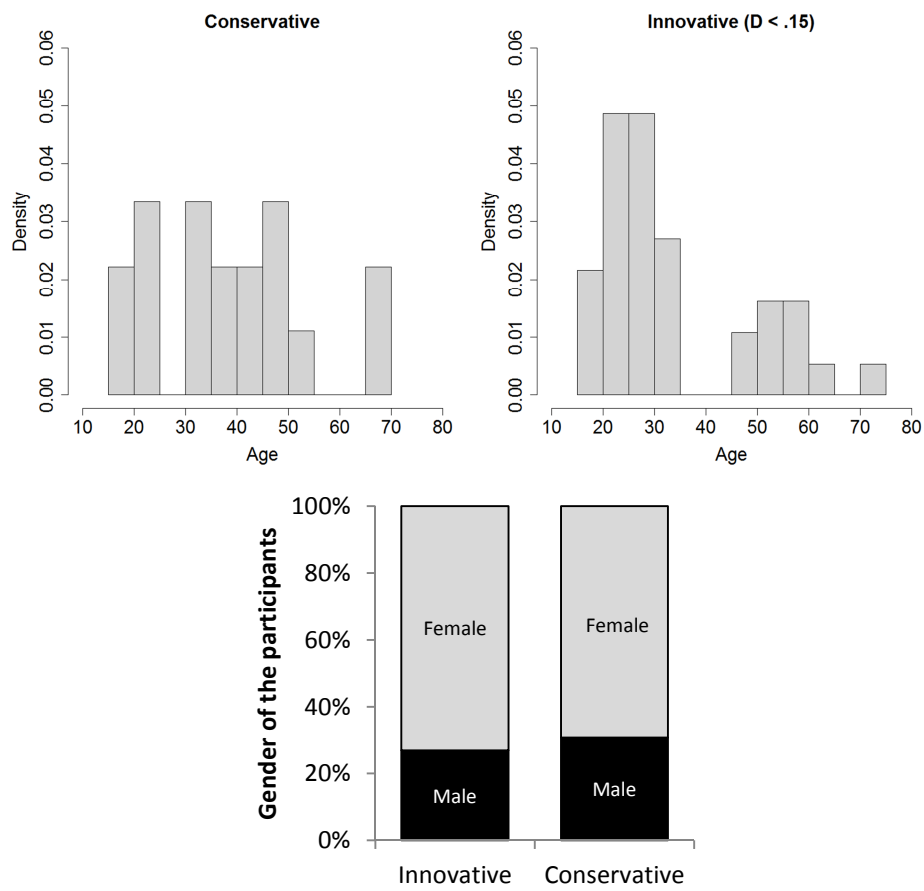
Figure 8: *Age and gender distribution of innovative and conservative speakers*

4.3 Conclusions

The statistical analysis of the obtained *D* measures proved that a randomly sampled group of informants provide data that follow normal distribution, i.e., we may conclude that the newly created LI-IAT appears to show no clear design-related bias.

The agreement between the measured implicit (*D* measure) and explicit (evaluative responses and self-report data) language attitude measures was not in line with some previous findings obtained by using IAT (see e.g., Pantos 2010, Pantos & Perkins 2013) or other techniques (Mády 2012), as in the present study a high ratio of agreement was found. This result might be attributed to several causes, but due to the small number of participants in the groups general conclusions should not be drawn. We venture to conclude, however, that there is a possibility that the extent to which the innovative form *"Lehet…"* is rejected among conservative users is smaller than the extent of the rejection of clearly and more commonly stigmatized (substandard or dialectal) language forms in the same group.[10] Nonetheless, it should be noted that the different types of explicit measurements provided highly different explicit attitude measures, and we should therefore also conclude that one should be cautious about the choice of measurement

---

[10] Such as, for instance, the errors, stigmatized forms and variants that are characteristic of spontaneous speech used in Mády (2012), or other variables that are generally said to be examples of stigmatization in Hungarian, e.g., *"innák már valamit"* instead of *"innék már valamit"* 'I really feel like drinking something', or *"majd meglássuk"* instead of *"majd meglátjuk"* 'we will see about that'.

technique and the interpretation of its data in explicit attitude detection. Based on the data on the relationship between explicit and implicit attitude we also see evidence for the common assumption that in most cases it is the innovative implicit attitude that is "denied" or covered explicitly (it is not accessible consciously for self-report, for instance), whereas conservative implicit attitude is in almost all cases paired with conservative explicit attitude.

Although the distribution of age in both the innovative and the conservative groups was in line with traditional sociolinguistic literature and it showed that an innovative attitude is more frequent in younger age groups, the gender distribution showed unexpected ratios, most probably due to the small number of participants in the resulting attitude groups and an unbalanced ratio of men and women participating in the study.

5   General Discussion

The literature of language change indicates that finding innovative speakers in the community is a key issue in the recognition of an ongoing language change. However, the identification of potentially innovative linguistic behavior is a challenging task. In the present paper we discussed the attitude construct and how this concept has been used in linguistic studies so far. We proposed that to solve the above mentioned problem, it is possible to expand the most common (or traditional) interpretation of language attitudes. This way, innovative speakers can be characterized by a specific implicit attitude towards linguistic innovations and linguistic change as such. For the purpose of operationalizing implicit attitude towards linguistic innovations, we proposed to adapt the auditory Implicit Association Test (IAT), a measurement tool of social psychology that uses auditory prompts, and includes a linguistic variable that is subject to an ongoing language change as a test variable (i.e., as target category). We argued that this new test, the LI-IAT, is capable of capturing a specific aspect of implicit attitude towards language change, and thus the LI-IAT measure may be an indicator of potentially innovative linguistic behavior. On this basis we also suggested that the LI-IAT is a tool that enables us to identify potentially innovative language users independently of their demographic characteristics. We proposed that, as a consequence, the LI-IAT may support the reliability of studies concerning the recognition of ongoing language change processes.

The method proposed in this paper, the LI-IAT was based on a test of Mády (2012) that used evaluative responses on substandard linguistic forms, speech errors and particular discourse markers characteristic of spontaneous speech to detect innovative and conservative implicit attitude. In the present study this idea was further developed in two aspects.

First, the efficacy of bypassing conscious control was increased by using a completely implicit task, by using the auditory IAT that uses no explicit evaluation at all, and where implicit evaluative responses are only inferred from reaction time data (i.e., IAT is a semi-online method). To emphasize the benefits of using IAT in implicit language attitude detection, it should be recalled that implicit measures are presumed to be relatively immune from many of the concerns that plague self-report measures. Therefore IAT is especially useful in domains in which social desirability is a concern, as e.g., in linguistic behaviour.

Second, the efficacy of the detection of innovators was increased, as we did not detect the respondents' susceptibility to identify speech errors, stable variations or

substandard forms (as in Mády 2012), but we detected the automatic reactions to an innovative linguistic form, a form that is assumed to be part of an ongoing language change. For this purpose an explicit attitude test was conducted in apparent time. Since it reflects beliefs about standards well, we take this explicit attitude test to be suitable to detect a linguistic variant that may be considered a linguistic innovation both from a diachronic and a synchronic point of view.

In the present study two experiments were reported. In Experiment 1 the linguistic variable was designated for the LI-IAT through an explicit test: 'presence or absence of the subordinator *hogy* in sentences with the verb *lehet*': e.g., "Lehet, hogy/ Lehet(,) *ő is eljön.*" "it's possible that/possible she comes over, too' (in short: "*Lehet, hogy…/Lehet*", conservative form/innovative form, respectively). The results of the study showed that younger adults accept the second, innovative variant of the variable "*Lehet…*" more than older adults, as the younger group accepted both competing variants in more than 40% of all cases, whereas older speakers accepted the "*Lehet…*" version in only about 10% of all cases. Hence, the phenomenon represented by the "*Lehet, hogy…/Lehet…*" example sentences was considered to be part of an ongoing language change.

The evaluative responses the participants gave on the test variants in Experiment 1 were also regarded as explicit attitude measures and used as such in the analysis of Experiment 2, in order to assess the interrelation between explicit and implicit attitude measures obtained in the present study. In another a pre-test (not reported here in detail) we, amongst others, made a first attempt to search for the sufficient category labels to be used in the LI-IAT to represent innovative and conservative linguistic variants. According to the results, the labels *older* and *newer* are not sufficient for the listeners to differentiate consistently between the two variants of the tested linguistic variables. Therefore, in the construction of the LI-IAT we opted for the use of direct metalinguistic labels similarly to the studies of Pantos (2010), Pantos and Perkins (2013), and Campbell-Kibler (2012).

In Experiment 2 the newly constructed LI-IAT was tested, and some simple explicit attitude measures (self-reports) were also recorded and contrasted with the implicit LI-IAT measures. The analysis consisted mainly of the basic exploration of the attitude construct captured by the LI-IAT. It was observed that the random sample of 40 participants provided *D* measures with normal distribution which is considered to be an indicator of a reliable data collection. Interestingly, in contrast with some previous findings (see e.g., Pantos 2010, Mády 2012, Pantos & Perkins 2013) the present study showed a high ratio of agreement between implicit and explicit attitude measures (the latter was assessed on the basis of explicit evaluation of example sentences from Experiment 1 and self-report from Experiment 2), while the two different kinds of explicit measures showed also a disagreement to a large extent (in 58%). Therefore, it was also concluded that the acceptability of the innovative "*Lehet…*" is most probably different from the explicitly stigmatized, stereotypical language forms, and that one should be careful when choosing a method for explicit attitude measurements. The age and gender distribution of innovative speakers and the age distribution of conservative speakers found in Experiment 2 met the expectations: innovators were mostly young adults, and females, whereas conservatives were found of almost all ages. Although the gender distribution of conservatives was somewhat unexpected, as innovators were found in both gender groups at almost the same percentage, this result may be due to the not well-balanced number of female and male participants in the study, thus from this finding no further conclusions was drawn.

It must be emphasized again that the proposed interpretation of innovative language attitude, that is, language attitude towards linguistic innovations and language change may not be understood as a direct predictor of innovative behavior in the production of the particular linguistic innovation at hand, as implicit attitudes are never taken as predictors of behavior without further experimental evidence gathered in psychology either. However, this interaction is actually of no great importance, if the measure of implicit attitude is interpreted merely as an indicator of tolerance or acceptance of linguistic innovations as such (and not towards the tested variable exclusively). This way the LI-IAT may be considered to be capable of operationalizing a specific aspect of implicit innovative linguistic attitude that may afterwards be also considered as a possible predictor of (perceptive or productive) innovative linguistic behavior in the case of other (independent) language change processes as an independent variable. In this sense, the LI-IAT measure reflects a particular aspect of implicit innovative linguistic attitude numerically, which is without a doubt a novel approach towards a research method that enables researchers to identify speakers of potentially innovative linguistic behavior.

It should be noted again that the optimal interpretation of the D measure, which is the most common numeric output of an IAT and our LI-IAT as well, has yet to be clarified. This will take with a significant amount of empirical data to be obtained in future research. To understand the nature of the implicit attitude captured by the LI-IAT we need to carry out several thorough empirical studies on language change processes that are already described for contemporary Hungarian where LI-IAT data may be confronted with linguistic data. One of our goals to pursue in the future is to investigate these interrelations, that is, to map the D measure obtained in the LI-IAT onto the innovative behavior observable in the speakers' speech production and perception.

As far as sound change processes are concerned, we plan to conduct a production experiment in which we obtain data by means of elicited speech (in which sound change processes concerning the Hungarian /iː uː yː/ are to be detected), and to redo a previously conducted perception experiment concerning the supposed change in the vowel length oppositions in Hungarian (Mády 2012) using the new LI-IAT for implicit attitude detection. Besides the obvious questions about the interrelation of implicit attitudes captured by the LI-IAT and linguistic data, we will also seek to identify the correct interpretation of the traditional effect size criteria in the present implementation of the IAT. In particular, we aim to answer two questions. First, where to draw the line between innovative and conservative attitudes and speakers. And second, whether the dimension covered by the D measures should be considered to be categorical or rather gradual.

In addition, in future research we wish to map the specific aspect of implicit innovative linguistic attitude captured by the LI-IAT to several types of language change processes (e.g., morphological change, syntactic change, and especially sound change processes) to assess which types of language change this specific aspect of implicit attitude (i.e., the D measure of the LI-IAT) is a reliable indicator of.

There is a lot of potential for using the IAT in studies in language attitude in particular and sociolinguistics in general. One question we also plan to investigate (in line with the ideas behind Campbell-Kilber 2012) is the meaning of stigmatization and the rejection of linguistic innovations by conservative speakers. With the LI-IAT at hand, it is now possible to understand what stigmatization or the rejection of certain linguistic forms means by testing possible hypotheses about their interpretation. In an IAT

paradigm, it is possible to align any type of concept pair. Thus, we may associate an innovative form, such as POSSIBLE and POSSIBLE THAT, with such notions as, for instance, *grammatical* versus *ungrammatical*, and with this type of alignment we actually test a commonly suggested hypothesis of sociolinguistics, namely that conservative speakers find innovative variants ungrammatical. To understand why this hypothesis may be of interest, consider, for instance, explicit attitude tests where it is always presupposed that non-standard forms are in a way "not correct", since stigmatization itself is measured on the basis of evaluation and grammaticality judgments that require participants to qualify test variables as *correct* and *incorrect*. Although this interpretation is often suggested (in some cases implicitly), there is yet no direct evidence for it. By adapting the IAT, however, we can take a step forward to find empirical support for or against this hypothesis.

With the LI-IAT it is also possible to associate innovative and conservative linguistic forms (again, e.g., POSSIBLE and POSSIBLE THAT) with such dimensions as *age* or the scales of *education, socioeconomic status*, etc., thus the stereotypical representations of innovative speakers is also more accessible for investigation than before.

Lastly, to raise further questions that are to be discussed in future research and that are to be kept in mind when using IAT in sociolinguistic research, two remarks from the literature of attitude and IAT measurements are cited from the field of social psychology. Karpinski and Hilton (2001) (as well as Greenwald and Banaji 1995) emphasize that implicit attitude is an introspectively undefined (or inaccurately identified) *trace of past experience* that mediates favorable or unfavorable thought, feeling, or action toward an object. On this basis the *environmental association model* claims that IAT is only able to measure the trace of past experience mentioned above, but we can have no certain knowledge of *how* the thought or action is mediated. Based on this theory, the association strength revealed by the IAT is only indicative of the case that the individual has been exposed to. This could be, for instance, a larger number of positive–*innovative form* and negative–*conservative form* associations than negative–*innovative form* and positive–*conservative form* associations. Karpinski and Hilton also claim the IAT may tell us what associations the person has been *exposed to* in his or her environment rather than the *extent* to which the person *endorses* the attitude. In addition they also claim that the environmental association model posits a dissociation between explicitly measured attitudes and the IAT consistent with Devine's (1989) dissociation of exposure to stereotypic knowledge (which may be measured by the IAT) and personal beliefs (which may be measured by explicit attitude scales). Another aspect to bear in mind is that there is evidence that only relatively spontaneous choices that do not involve a great deal of personal involvement are assumed to be influenced by implicit attitude (Karpinsi & Hilton 2001, cf. Dovidio et al. 1997): attitudes that are measured implicitly tend to predict spontaneous or nonverbal behaviors (blinking, amount of eye-contact), whereas those measured explicitly tend to predict deliberative behaviors (e.g., consumer behavior).

In the present study an IAT measuring implicit attitudes towards linguistic innovations and language change was developed, and we concluded that it is a promising tool in sociolinguistic research. However, it is also intended to remind all the future users of this method (or any other research concerned in the field of attitude) to keep the warnings of social psychology in mind on the yet, to say, still "mysterious" nature of attitude, on the non-trivial interrelations between implicit attitude and action, and on the concerns of attitude detection.

References

Ajzen, Ice K. 2001. Nature and operation of attitudes. *Annual Review of Psychology* 52. 27–58.

Bácskai-Atkári, Júlia & Dékány, Éva. 2014. From non-finite to finite subordination: The history of embedded clauses. In Katalin É. Kiss (ed.), *The Evolution of Functional Left Peripheries in Hungarian Syntax*. 148–223. Oxford: Oxford University Press.

Cacioppo, John T., Stephen L. Crites Jr., & Wendi L. Gardner. 1996. Attitudes to the Right: Evaluative Processing is Associated with Lateralized Late Positive Event-Related Brain Potentials. *Personality and Social Psychology Bulletin* 22(12). 1205–1219.

Campbell-Kibler, Kathryn. 2012. The Implicit Association Test and sociolinguistic meaning. *Lingua* 122. 753–763.

Croft, William. 2000. *Explaining Language Change. An Evolutionary Approach*. (Longman Linguistics Library) London: Longman.

Devine, Patricia G. 1989. Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology* 56(1). 5–18.

Dovidio, John F., Kawakami, Kerry, Johnson, Craig, Johnson, Brenda & Howard, Adaiah. 1997. On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology* 33(5). 510–540.

E. Abaffy, Erzsébet. 1976. Valószínűleg, hogy…? *Magyar Nyelvőr* 100. 397–398.

É. Kiss, Katalin. 2010. Valószínűleg, hogy román kontaktushatás. In Katalin É. Kiss & Attila Hegedűs (eds.), *Nyelvelmélet és kontaktológia*. 223–238. Piliscsaba: PPKE BTK.

Egedi, Barbara. 2014. The DP-cycle in Hungarian and the functional extension of the noun phrase. In: Katalin É. Kiss (ed.), *The Evolution of Functional Left Peripheries in Hungarian Syntax*. 56–82. Oxford: Oxford University Press.

Frijda, Nico H. 1986. *The Emotions*. Cambridge UK: Cambridge University Press.

Gawronksi, Bertram & de Houwer, Jan. 2014. Implicit measures in Social and Personality Psychology. In Harry T. Reis & Charles M. Judd (eds.), *Handbook of Research Methods in Social and Personality Psychology* (2nd edition). 283–310. New York: Cambridge University Press.

Givón, Talmy. 2001. *Syntax. An Introduction. Volume 1*. Amsterdam: John Benjamins.

Grétsy, László & Miklós, Kovalovszky (eds.). 1980. *Nyelvművelő kézikönyv*. Budapest: Akadémiai Kiadó.

Greenwald, Anthony G., & Banaji, Mazarin R. 1995. Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological Review* 102. 4-27.

Greenwald, Anthony.G., McGhee, Debbie E. & Schwartz, Jordan L.K. 1998. Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology* 74. 1464–1480.

Greenwald, Anthony G., Nosek, Brian A. & Banaji, Mazarin R. 2003. Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology* 85(2). 197–216.

Haader Lea. 1991. Az összetett mondatok. In Loránd Benkő (ed.), *A magyar nyelv történeti nyelvtana I*. 728–741. Budapest: Akadémiai Kiadó.

Haader Lea. 1995. Az összetett mondatok. In Loránd Benkő (ed.), *A magyar nyelv történeti nyelvtana II/2*. 506–665. Budapest: Akadémiai Kiadó.

Haader Lea. 2001. Mikrodiakrónia és változásvizsgálat. *Magyar Nyelvőr* 125. 354–371.

Harris, Alice C. & Campbell, Lyle. 1995. *Historical syntax in cross linguistic perspective*. Cambridge: Cambridge University Press.

Heine, Bernd & Kuteva, Tania. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

Holtgraves, Thomas. 2004. Social desirability and self-reports: testing models of socially desirable responding. *Personality and Social Psychology Bulletin* 30(2). 161–172.

Horwitz, Suzanne R. & Dovidio, John F. 2015. The rich – love them or hate them? Divergent implicit and explicit attitudes toward the wealthy. *Group Processes and Intergroup Relations*. 1–29.

Hopkins, A. Nicholas. 1977. Historical and Sociocultural Aspects of the Distribution of Linguistic Variants in Highland Chiapas, Mexico. In Ben G. Blount & Mary Sanches (eds.), *Sociocultural Dimensions of Language Change*. 185–225. New York: Academic Press.

I. Gallasy, Magdolna. 1991. A névelő és a névelő-előzmény. In Loránd Benkő (ed.), *A magyar nyelv történeti nyelvtana I*. 461–475. Budapest: Akadémiai Kiadó.

I. Gallasy, Magdolna. 1992. A névelők. In Loránd Benkő (ed.), *A magyar nyelv történeti nyelvtana II/1*. 716–771. Budapest: Akadémiai Kiadó.

Karpinski, Andrew & Hilton, James L. 2001. Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology* 81(5). 774–788.

Kenesei, István. 1992. Az alárendelő mondatok szerkezete. In Ferenc Kiefer (ed), *Strukturális magyar nyelvtan 1. Mondattan*. 529–714. Budapest: Akadémiai Kiadó.

Kiefer, Ferenc. 2000. A szóösszetétel. . In Ferenc Kiefer (ed), *Strukturális magyar nyelvtan 3. Morfológia*. 519–568. Budapest: Akadémiai Kiadó.

Kiss, Jenő. 2000. Magyar nyelvjárási beszélők anyanyelvi tetszési indexe. *Magyar Nyelvőr* 124. 465–468.

Karpinski, Andrew & Hilton, James L. 2001. Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology* 81(5). 774–788.

Korompay, Klára. 1991. A névszójelezés. A névszóragozás. In *A magyar nyelv történeti nyelvtana I*. 259–318. Budapest. Akadémiai Kiadó.

Korompay, Klára. 1992. A névszójelezés. A névelők. In Benkő, Loránd (ed.), *A magyar nyelv történeti nyelvtana II/1*. 321–410. Budapest. Akadémiai Kiadó.

Labov, William. 1963. The social motivation of a sound change. *Word* 19. 273–309. [Revised as Ch.1 of Labov, William 1972. *Sociolinguistic Patterns*. Philadelphia, University of Pennsylvania Press, 1972. 1-42.]

Labov, William. 1966 [2006]. *The social stratification of English in New York City*. Cambridge: Cambridge University Press.

Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Labov, William. I980. *Locating language in time and space*. New York: Academic Press.

Lambert, W.E., R.C. Hodgson, R.C. Gardner, & S. Fillenbaum. 1960. Evaluational reactions to spoken language. *Journal of Abnormal and Social Psychology 60*.1: 44-51.

MacDonald Tara K. & Zanna Mark P. 1998. Cross-dimension ambivalence toward social groups: can ambivalence affect intentions to hire feminists? *Personality and Social Psychology Bulletin* 24(4). 427–441.

MacFarlane, A.E. & Stuart-Smith, J. 2012. 'One of them sounds sort of Glasgow Uni-ish'. Social Judgements and fine phonetic variation in Glasgow. *Lingua* 122(7). 764–778.

Mády, Katalin. 2010. Hungarian vowel quantity neutralization as a potential social marker. *Acta Linguistica Hungarica* 57(2–3). 167–188.

Mády, Katalin. 2102. How can a conservative language user be identified? – an approach via implicit and explicit language attitude. *2nd Workshop on Sound Change*. Kloster

Seeon, Bavaria, Germany, 2012. 2–4 May. http://www.phonetik.uni-muenchen.de/institut/veranstaltungen/soundchange/abstracts/mady.pdf

Mády, Katalin & Rácz, Péter. 2013. Attitude as a dimension of sound change. *Proceedings of the 6th Phonetics and Phonology in Iberia*. Lisbon, 24. July 2013. http://ww3.fl.ul.pt/laboratoriofonetica/papi2013/files/posters/poster128.html

Maniaci, Michsel R. & Rogge, Donald D. 2014. Conducting research on the internet. In Harry T. Reis & Charles M. Judd (eds.), *Handbook of Research Methods in Social and Personality Psychology* (2nd edition). 443–470

Milroy, James & Milroy, Lesley. 1985. Linguistic change, social network and speaker innovation. *Journal of Linguistics* 21(2). 339–384.

Nádasdy, Ádám. 2008. *Prédikál és szónokol.* Budapest: Magvető.

Pantos, Andrew J. 2010. *Measuring implicit and explicit attitudes toward foreign-accented speech.* : Texas: Rice University dissertation.

Pantos, A. J., & Perkins, A. W. 2013. Measuring implicit and explicit attitudes toward foreign accented speech. *Journal of Language and Social Psychology 32*(1), 3-20.

Ohala, John J. 1981. The listener as a source of sound change. In C. S. Masek, R. A. Hendrick, & M. F. Miller (eds.), *Papers from the Parasession on Language and Behavior.* 178–203. Chicago: Chicago Linguistic Society.

Oravecz, Csaba, Tamás, Váradi & Bálint, Sass. 2014. The Hungarian Gigaword Corpus. In: *Proceedings of LREC 2014.* http://www.lrec-conf.org/proceedings/lrec2014/pdf/681_Paper.pdf

Paulhus, Delroy L. 1991. Measurement and control of response bias. In J. P. Robinson, P. R. Slaver & L. S. Wrightsman (eds.), *Measures of personality and social psychological attitudes.* 17–59. San Diego: Academic Press.

Preston, Dennis R. 2003. Language with an Attitude. In J. K. Chambers, P. Trudgill & N. Schilling-Estes (eds.), *Handbook of Language Variaton and Change.* 40–66. Oxford: Blackwell.

R Core Team 2013. *R: A Language and Environment for Statistical Computing.* Computer program. Vienna, Austria. http://www.R-project.org

Sándor, Klára, Juliet Langman & Pléh Csaba. 1998. Egy magyarországi „ügynökvizsgálat" tanulságai: A nyelvváltozatok hatása a személypercepcióra. *Valóság* 8. 29–40.

Sankoff, Gillian. 2006. Age: Apparent Time and Real Time. In Keith Brown (editor-in-chief), *Elsevier Encyclopedia of Language and Linguistics* Vol. 1 (2nd edition). 110–116. Elsevier: Amsterdam.

Szabolcsi, Anna & Tibor Laczkó. 1992. A főnévi csoport szerkezete. In Ferenc Kiefer (ed), *Strukturális magyar nyelvtan 1. Mondattan.* 179–298. Budapest. Akadémiai Kiadó.

Vande Kamp, Mark E. 2002. *Auditory Implicit Association Tests.* Washington: University of Washington dissertation.

Virtanen, Susanna 2013. Contextual function of noun marking in the direct object–marking system of Eastern Mansi. In Márta Csepregi, Kata Kubínyi & Sivonen Jari (eds.), *Grammatika és kontextus: új szempontok az uráli nyelvek kutatásában III.* Uralisztikai Tanulmányok 20. 305–317. Budapest: ELTE Finnugor Tanszék.

Wagner, Suzanne Evans. 2012. Age grading in sociolinguistic theory. *Language and Linguistics Compass* 6(6). 371–382.

Wilson, Timothy D., Lindsey, Samuel & Schooler, Tonya Y. 2000. A model of dual attitudes. *Psychological Review* 107(1). 101–126.

Andrea Deme
MTA Nyelvtudományi Intézet
deme.andrea@nytud.mta.hu

Katalin Gugán
MTA Nyelvtudományi Intézet
gugan.katalin@nytud.mta.hu

Bálint Sass
MTA Nyelvtudományi Intézet
sass.balint@nytud.mta.hu

Katalin Mády
MTA Nyelvtudományi Intézet
mady.katalin@nytud.mta.hu

The Mansi Ditransitive Constructions

Bernadett Bíró – Katalin Sipőcz

The aim of this paper is to investigate Mansi ditransitive constructions from a typological point of view. Mansi has an alternation of indirective (indirect object) and secundative (secondary object) ditransitive constructions. Also passivization plays an important role in alternation. In Mansi both constructions can passivize, although the passivization of the secundative construction is more frequent. The alternation is related to topicality, the choice between the two ditransitive constructions (and the conjugation types and voice of the verb) is made in order to express the relative topicality of the arguments.

Keywords: *Mansi language, ditransitive constructions, ditransitive alternation*

1   Introduction

The aim of this paper is to present Mansi ditransitive constructions and to describe the universal and specific properties of transitivity alternations in Mansi. The investigation is put into a typological frame and the findings are related to typological classifications presented by Malchukov, Haspelmath and Comrie (2010).

The Mansi ditransitive constructions have been widely investigated in Uralic studies, but the term "ditransitive" has been connected to the phenomenon only recently. A characteristic feature of Ob-Ugric languages is the two ways of expressing ditransitive situations, see (1) and (2). Traditionally this has been considered an exotic feature of the Ob-Ugric languages, unknown in other related languages. In the last decades, these constructions were put in a new perspective by studies on the transitivity of Ob-Ugric languages (Nikolaeva 2001, Skribnik 2001, Szilágyi 2014, Virtanen 2011, 2013). The discussion of these constructions in a typological frame has not been done yet, the authors of the present study are the first to examine this syntactic feature of the Mansi language from a typological point of view (Bíró 2013, 2015, Sipőcz 2011, 2013, 2015a, 2015b, 2016).

(1) a. *am   naŋən(n) sūp   junt-ey-əm*   (Mansi)
        I   you.LAT   shirt   sew-PRS-1SG
     b. *am   naŋən   sūp-əl   junt-i-ləm*
        I   you.ACC   shirt-INSTR   sew-PRS-SG.1SG[1]
        'I sew a/the shirt to you.'

(2) a. *(ma) Juwan-a   ān   ma-s-əm*   (Northern Khanty)
        I   John-LAT   cup   give-PST-1SG
     b. *(ma) Juwan   ān-na   ma-s-ēm*
        I   John   cup-LOC/INSTR   give-PST-SG.1SG
        'I gave a/the cup to John.'

                                                            (Nikolaeva 2001: 32)

---

[1] We gloss the personal suffixes in the following way: first we indicate the number of the object/possession, then the person and number of the subject/possessor. E.g. DU.3SG: (1) a verb suffix agreeing with the object in dual and 3rd person singular subject; or (2) a possessive suffix agreeing with the possessum in dual and 3rd person singular possessor. In subject agreement the gloss indicates the number and person of the subject. The singularity of the possessum is not indicated according to its unmarkedness.

The paper is organized as follows: The short description of the Mansi language (1.1) and the corpus (1.2) is followed by the typological characterization of the ditransitive constructions (2.1). Then we introduce the Mansi constructions and their use (2.2). In the next parts we analyze the alternation (3) and the passivization (4) of Mansi ditransitive constructions. Finally, Section (5) concludes with a summary and statistics of our findings.

1.1 The Mansi language

The Mansi (or Vogul) language is one of the most endangered languages of the Uralic language family. Even at the beginning of the 20th century, four dialects of Mansi were still spoken (Northern Mansi, Eastern Mansi, Western Mansi and Southern Mansi), but today only the Northern Mansi dialect is used, spoken by less than 1000 people (See Table 1). Northern Mansi is currently threatened by the process of language shift to Russian, almost all of its speakers are bilingual. Nowadays, the term Mansi is usually used as referring to the Northern Mansi dialect. In this study we concentrate only on the ditransitive constructions of Northern Mansi, all of our examples are from this dialect.

MANSI	LANGUAGE PROFICIENCY
12,269	938 (7.6%)

Table 1. *Ethnically Mansi population and language proficiency according to census data 2010[2]*

1.2 The corpus

The corpus used for this investigation consists of representatively selected 200 clauses containing ditransitive constructions. The language data are taken from a large number of written sources and extend over about a hundred years. They have been collected from older and current collections of folkloric texts (VNGy I-IV, Kálmán 1976), collections of interviews with Mansi individuals (Dinislamova 2007), and newspaper texts (LS). Certain ditransitive verbs and constructions are very common (e.g. 'somebody gives something to someone', 'somebody tells, says something to someone', etc.), these frequent examples were selected. We tried to compile our corpus so that it would include as many verbs used in several constructions and contexts as possible. In addition to this, full structures (i.e. structures with more arguments) were preferred. Although language use in the older and newer collections show significant differences as regards not only vocabulary but also many aspects of grammar (Bíró & Sipőcz 2009), changes in the ditransitive contructions in this respect is not discussed in this paper.

2 Ditransitive constructions

2.1 The typology of the ditransitive constructions

In linguistic typology, ditransitive constructions have recently become a popular topic of international research, several studies and books focus on ditransitivity (e.g. Malchukov et al. 2010). Smaller Uralic languages, however, are entirely missing from these studies.

A ditransitive construction consists of a ditransitive verb, an agent argument (A), a recipient (recipient-like, addressee) argument (R) and a theme argument (T) (Malchukov et al. 2010, 1):

---

[2]   http://www.perepis-2010.ru/results_of_the_census/results-inform.php

(2)   *Mary gave   John a book. / Mary told   John a story.*
      A   R   T   A   R   T

Ditransitive verbs are typically physical transfer verbs such as *give, send, sell, bring*, etc., but in most languages some verbs expressing mental transfers (verbs of communication) like *say, tell, show*, etc., also behave syntactically in a similar way. In some languages also the benefactive verbs like *make, cook, build*, etc. and some caused motion verbs like *throw, drop*, etc. appear in the same construction and in this case these are also listed as ditransitive verbs. The group of those verbs which appear in ditransitive constructions is particularly extended in Mansi, all of the above mentioned group of verbs take the same argument structure (Sipőcz 2015b, 2016).

The most general typological characterisation of ditransitive constructions is based on the comparison of monotransitive and ditransitive constructions. The basic alignment types are distinguished on the basis of the encoding of the T (theme) and the R (recipient) compared to the patient (P) in the monotransitive construction (Malchukov et al. 2010, 3). According to this there are three basic types:

1. the *indirect object construction* (IOC) (or *indirective alignment*), where the T and the P have the same morphological marking (it can be zero as well), but the R is treated differently from the T and P, e.g.

(4)   (monotransitive)   *János könyv-et   olvas.*   (Hungarian)
                                  János book-ACC   read.3SG
                                  'János reads a book.'

      (ditransitive)   *János könyvet   vesz   Anná-nak.*
                                  János book-ACC   buy.3SG Anna-DAT
                                  'János buys a book to Anna.'

2. the *secondary object construction* (SOC) (or *secundative alignment*), where the P and the R have the same marking and the T is treated differently. (It is also called a primary object construction.)

(5)   (monotransitive)   *Uukaraawiciizɨ tɨɨri   me-wa-zeiya.*   (Huichol, Mexico)[3]
                                  women   children 3PL.NOM-3PL.PRIM-see
                                  'The women see the children.'
      (ditransitive)   *Nee   tumiini   uukari   ne-wa-ruzeiyastɨa.*
                                  I   money   girls   1SG.NOM-3PL.PRIM-show
                                  'I showed the money to the girls.'
                                                (Haspelmath 2005)

3. the *double object construction* (DOC) (or *neutral alignment*), where the P, the R and the T are encoded in the same way. This type is well-known from English, cf.:

(6)   a. *Mary saw **John**.*   (English)
      b. *He gave **John** a **book**.*

---

[3]   Comparing examples (4) and (5) we can see that the marking of the arguments can appear on the noun (flagging: (4)) or on the verb (indexing: (5)) or on both.

Further types which are logically also possible but can be disregarded due to their minimal occurrence, are the so-called tripartitive (T ≠ R ≠ P) and horisontal (T = R ≠ P) constructions (Malchukov et al. 2010, 5–6).

Finally, we have to mention that there are two further kinds of ditransitive constructions that are impossible to fit into the aforesaid classification. These types are not based on the comparison of monotransitive and ditransitive clauses, the indirective and secundative characters are however clearly distinguishable in their cases, too. These are the serial verb construction and the possessive construction. (Malchukov et al. 2010, 11–15, Margetts & Austin 2007) In serial verb construction the marker of the T or the R is a grammaticalized verb, see the Yoruba example (7) in which the verb 'give' is the marker of the R argument.

(7) *Ó tà-á   fún   mi*   (Yoruba)
     he sell-it give me
     (a) 'He sold it to me.' (b) 'He sold it for me.'

(Margetts & Austin 2007)

In the possessive construction, the R appears as the possessive modifier of the T, this can be exemplified from Nganasan where the R argument appears as the possessor of the T argument.

(8) *mənə kńiga-ðə-mtu   mi-śiə-m*   (Nganasan)
     I   book-DST-ACC.3SG   give-PST-1SG
     'I gave him/her the book.'

(Wagner-Nagy & Szeverényi 2013, 28)

2.2 The Mansi ditransitive constructions

Mansi belongs to languages having more than one ditransitive constructions. The Mansi constructions are:

(I.) *Indirect object construction,* where the theme (T) of the ditransitive construction is the syntactic object, and the recipient (R) is encoded with the lative-dative *-n* suffix (LAT). The verb can be in the subjective (9) or objective conjugation (10).[4] There is no accusative case in Northern Mansi except for the personals pronouns, which have a distinct accusative form. (It is very unusual, however, that in IOC the T is a personal pronoun. There is no example for this in our corpus. T expressed by a personal pronoun appears typically as the subject of passive sentences, cf. example (28).)[5]

(9) *Pjotr Gavrilovič ānəmn   jurt-ane   jot   t'it   kassēta-γ   t'ēt-əs.*
     P.G.   I.LAT   friend-PL.3SG with two   cassette-DU send-PST.3SG
     'Pjotr Gavrilovich sent me two cassettes with his friends.'

(Dinislamova 2007, 5)

---

[4] These two conjugation types, characteristic of some other Uralic languages as well, have been referred to also as *indeterminate* and *determinate conjugations, indefinite* and *definite conjugations* as well as *subject-verb agreement* and *object-verb agreement* (Kálmán 1976, Honti 1988, Keresztes 1998). In this study we use the terms *subject-verb agreement* and *object-verb agreement* as well as *subjective* and *objective conjugations,* since when the former conjugation type is used it means that the verb agrees only with the subject, while in the case of the latter the verb agrees with both the subject and the direct object.

[5] In Mansi, which has a rather strict SOV word order, IOCs containing a verb agreeing only with the subject have an Agent - Recipient - Theme - Verb order in most cases, while IOCs containing a verb agreeing with both the subject and the direct object have an A - T - R - V order in most cases. (Cf. Bíró 2015)

(10) *akw' sup-ä   kat'i-tä-n   mis-tä,*
      one   piece-3SG   cat-3SG-LAT   give.PST-SG.3SG
      akw' sup-ä   āmp-ən   mis-tä.
      one   piece-3SG   dog-LAT   give.PST-SG.3SG
      '(S)he gave one piece to his/her cat and the other one to the dog.'

                                                                    (VNGy IV, 343)

(II.) *Secondary object construction,* where the R of the ditransitive construction is an unmarked object and the T is marked with the instrumental -*l* suffix (11). If R is a personal pronoun, it is in ACC form (12). In this construction, the verb is almost always in the objective conjugation (in more detail see section 3).

(11) *Mań piγ-əm   nē-γəl   viγ-ləm.*
      little son-1SG   woman-INSTR   take.PRS-SG.1SG
      'I will find a wife for my youngest son.'

                                                                    (VNGY IV, 324)

(12) *Nēnan   am   śopr-śonaχ-əl   wār-i-jayəm.*
      you(DU).ACC   I   silver-cup-INSTR   make-PRS-DU.1SG
      'I make the two of you a silver cup.'

                                                                    (Kálmán 1976, 70)

3   Alternation

Several languages have more than one ditransitive constructions. This phenomenon is called alternation, and is well-known from English, e. g.:

(13) a. *Mary gave a pen to John.*
      b. *Mary gave John a pen.*

In English, the indirective and the neutral alignments alternate.[6] In Mansi, we can see the alternation of indirective and secudative types. This latter type of alternation is cross-linguistically more common than the alternation found in English. The alternation of indirective and secudative constructions can be found also in European languages, but it is usually limited to only a few verbs (cf. Malchukov et al. 2010, 18–19). E.g.:

(14) IOC: *Sütemény-t   és   kávé-t   kínál-ok   a   vendég-ek-nek.*   (Hungarian)
           cake-ACC   and   coffee-ACC   offer-1SG   ART   guest-PL-DAT

      SOC: *Sütemény-nyel   és   kávé-val   kínál-om   a   vendég-ek-et.*
           cake-INSTR   and   coffee-INSTR offer-1SG.O   ART   guest-PL-ACC
           'I am offering cookies and coffee for the guests.'[7]

---

⁶   The phenomenon is also called "Dative Shift" and there is an enormous literature on this subject (Givon 1984, Beck & Johnson 2004, Rappaport & Levin 2008 among others).

⁷   The number of those verbs which can appear in both IOC and SOC constructions is strongly limited also in Hungarian, and usually even these verbs show differences concerning the use of the preverbs, e.g. *(meg)kínál vkit vmivel* (SOC) *vs. kínál vkinek vmit* (IOC) 'offer sth to sb to drink/eat', *megajándékoz vkit vmivel* (SOC) vs. *ajándékoz vkinek vmit* (IOC) 'give a gift'.

Mansi, however, seems special typologically regarding that there is no restrictions in the use of the verbs: almost any ditransitive verb can appear in both constructions. (Cf. Sipőcz 2015b, 2016)

Concerning alternations the important question is which factors determine the choice between the different constructions. According to the typological studies several factors can be mentioned: the markedness of the arguments, the prominence differences between the T and R arguments (e.g. in animacy, nominal/pronominal status, discourse status (topicality) cf. Bresnan et. al 2007), there may be semantic difference between the alternating constructions, etc. It often occurs that the choice between the different constructions is simultaneously affected by more than one factor in a given language. In English for example, the Double Object Construction is favoured when R is more prominent than T (e.g. R is a pronoun), and/or when R is more topical than T (Malchukov et al. 2010, 20–21).

As it was already mentioned, in some languages the alternation is related to topicality. Ob-Ugric languages seem to belong to this group. The use of subjective and objective conjugations in Mansi has been discussed in detail for decades (cf. for example Lavotha 1953, Honti 1969, Skribnik 2001, Dolovai 2003) and in these studies often also the alternation of the two ditransitive constructions has been examined from a historical, morphosyntactic and functional grammatical point of view. (However, the term "ditransitive construction" was not used in these studies.) There are diverse statements about the Ob-Ugric alternation in the literature. Concerning Khanty, it has been suggested that the alienability or inalienability of the object (theme) (Honti 1999) and/or the focal status of the object (id.) influences the choice. According to Rombandeeva (1979, 99–115), the only native Mansi linguist, the choice is affected by the definiteness and by the fact how emphasized the theme is. Kulonen discusses these constructions in connection with Dative Shift and she claims that the aim of switching from one construction to the other is to promote the Recipient to direct object position, from where it could also be promoted to subject position with the help of passivization. (Kulonen 1999)

The connection between the use of different conjugations, constructions and topicality was studied by Nikolaeva in Khanty (2001), and by Skribnik in Mansi (2001). According to Skribnik (2001), the function of promoting the Recipient to direct object position is to express the relative topicality of different noun phrases within a clause. (She uses the term Dative Shift, similarly to Kulonen.) Thus the alternating constructions put either T or R into the position of the direct object, and the topicality of the direct object is indicated by the objective conjugation of the verb.

(15)   a. IOC: there is no topical object – the subjective conjugation is used
        Am   tawen   mōjt   mōjt-eɣ-um.
        I   (s)he.LAT   tale   tell-PRS-1SG
        'I tell him a tale.' (< 'What do you do?')

    b. IOC: the T is topical – the objective conjugation is used
        Am   mōjt   tawen   mojt-i-lum.
        I   tale   (s)he.LAT   tell-PRS-SG.1SG
        'I tell him the tale.' (< 'Who do you tell the tale to?')

    c. SOC: the R is topical – the objective conjugation is used
        Am   tawe   mōjt-əl   mōjt-i-lum.
        I   he.ACC   tale-INSTR   tell-PRS-SG.1SG
        'I tell him a tale.' (< 'What do you tell him?')

                                        (Skribnik 2001, 228)

Our data generally supports this claim. Sentences in (16) are taken from the beginning of a Mansi tale, the sentences are quoted in the original order (as found in the tale). In the first sentence, there is no topical[8] direct object thus the indirective construction is used. In the second sentence, both the main hero and the arrow with the bow are actually topical. The choice of the secundative construction is motivated by the fact that the main hero is pragmatically more important as the primary topic of the tale, therefore it is put to the direct object position and indicated by the objective conjugation of the verb. The third sentence is about the main hero, who takes the role of the Recipient, the Agent is not important thus passivization is used.

(16) a. *ākəmēkw,   ānəmn   ńāl   wār-n,   jōwt   wār-n*
          aunt   I.LAT   arrow   make-IMP.2SG   bow   make-IMP.2SG
          'Auntie, make me an arrow, make me a bow.'

      b. *ńāl-l   wār-i-ləm,   jōwt-əl   wār-i-ləm*
          arrow-INSTR   make-PRS-SG.1SG   bow-INSTR   make-PRS-SG.1SG
          'I provide you with an arrow, I provide you with a bow.'

      c. *ja-ti, ńāl-l   wār-we-s,   jōwt-əl   wār-we-s*
          well   arrow-INSTR   make-PASS-PST.3SG   bow-INSTR   make-PASS-PST.3SG
          '(S)he was provided with an arrow, (s)he was provided with a bow.'

                                                                    (Kálmán 1976, 68)

There are also some examples (17) where this kind of connection between topicality and the chosen ditransitive construction cannot be stated this clearly. In the corpus, there are several examples in which the use of the given construction is hard to explain. The sentences of example (17) were uttered in similar situations, the constructions were still different.

(17) a. *Ānəm   tē-ne   matər   tot-en,   sim-əm*
          I.LAT[9] eat-PTCP.PRS something bring-IMP.2SG   heart-1SG
          ētχəl-awe!
          starve-PASS.PRS.3SG
          'Give me something to eat, my heart is starving.'

      b. *Ānəm   tēn-ut-əl   tot-eln,   sim-əm*
          I.ACC   eat-thing-INSTR   bring-SG.IMP.2SG heart-1SG
          ētχəl-awe!
          starve-PASS.PRS.3SG
          'Give me something to eat, my heart is starving.'

                                                                    (VNGy I, 11)

In Table 3 (Section 5.1.), we can see the quantitative data representing the alternation of the ditransitive constructions without the passive ones. In our corpus, the IOC construction prevails. Regarding IOCs, there are twice as many examples with the verb in the subjective conjugation (the subject is the only topic), while approximately one third of these constructions contain a verb in the objective conjugation (both the subject and T are topical). In the SOC constructions, usually the objective conjugation is used. (Beside the subject, the R is also topical.) (Cf.: Tables 4 and 5) In this construction the direct object (R) is often omitted, cf. (18), (19).

---

[8]   By topic(al) we mean a previously mentioned or situationally given information (cf. Dalrymple & Nikolaeva 2014, 48–57).

[9]   The Lative/Dative suffix *-n* is often ommitted when the stem ends in a nasal consonant. (See example 1.)

(18) *Mir-əl   śōpiγt-i-ləm,   karapli-l   śōpiγt-i-ləm.*
     people-INSTR   equip-PRS-SG.1SG ship-INSTR equip-PRS-SG.1SG
     'I equip you with people, I equip you with a ship.'

(VNGy IV, 341)

(19) *Ńāl-əl   liγ-aγmēn.*
     arrow-INSTR   shoot.PRS-DU.1DU
     'We (two) shoot an arrow toward the two of you.'

(Kálmán 1976, 64)

As it was mentioned before, the verb in the SOC is usually in the objective conjugation. This clearly follows from the fact that the SOC is used in case of a topical Recipient (taking the position of the direct object) and a topical direct object is accompanied by a verb in the objective conjugation. Nevertheless, there are examples in our corpus in which the secundative constructions contain a verb in the subjective conjugation. In each sentence the subject is 2nd (20) or 3rd (21) person while the direct object (R) is 1st person, so it stands higher in the hierarchy of prominence (1SG > 1PL > 2SG > 2PL > 3SG > 3PL).

(20) *sis=jōr-əl=ke*   *naŋ*   *ānəm*   *pin-eγ-ən,*
     back-strength-INSTR=if   you   I.ACC   put-PRS-2SG
     maγil=jōr-əl=ke   *naŋ*   *ānəm*   *pin-eγ-ən*
     chest-strength-INSTR=if you   I.ACC   put-PRS-2SG
     'if you provide me with back-strength, if you provide me with chest-strength'

(VNGy II, 142)

(21) *jäγ-əm*   *sēl-əm*   *ōln-nəl*   *akw*   *ōln-pāl-əl*
     father-1SG   gather-PTCP.PST   money-ABL   one   money-half-INSTR
     ānəm   *at*   *majl-əs.*[10]
     I.ACC   NEG   give-PST.3SG
     'He did not give me even a half penny from the wealth gathered by my father.'

(VNGy IV, 343)

In these sentences the animacy hierarchy can explain the use of the subjective conjugation. In some Uralic languages using objective conjugation and similarly in some Eastern-Siberian languages (e.g. Chukchi, Koryak, Kamchadal), the so-called Inverse Agreement Constraint can be observed. This constraint means that if the object is more prominent in the animacy hierarchy than the subject, the verb agrees only with the subject (É. Kiss 2010, 2013). Thus in Hungarian, in Eastern Khanty and in the Samoyedic languages, the objective conjugation is used only if the object is 3rd person (Dalrymple & Nikolaeva 2011).

Mansi, however, seems irregular in this respect since the use of the objective conjugation is independent from the person of the object (Keresztes 1998, 417, Kulonen 2007, 111, Dalrymple & Nikolaeva 2011, 196). Numerous examples support this claim, e.g.:

(22) *ja,*   *at*   *pūw-i-te-e!*
     yes   NEG   catch-PRS-SG.3SG-EMPH
     'Yes, (s)he will not catch me!'

(Kálmán 1976, 144)

---

[10] It must be noted that the form *majləs* is irregular.

(23) *Manriɣ   at   wiɣ-lən?*
      why   NEG   take.PRS-SG.2SG
      'Why do you not take me?'

(24) *Taw   naŋən   ērupt-i-te.*
      (s)he   you.ACC   love-PRS-SG.3SG
      '(S)he loves you.'

On the basis of these examples, we can claim that the Inverse Agreement Constraint does not work in Mansi. It is noteworthy, however, that our irregular examples contain an R in the first person and an A in the second and third persons.

4   Passivization of ditransitive constructions

Beside the active ditransitive constructions, our corpus contains a great number of passive ditransitive constructions as well (see Table 2, Section 5.1.).

Concerning the passivization of ditransitive verbs, the question is which argument (T, R) can passivize (i.e. can be promoted to subject position). On the basis of this, three primary alignment types can be distinguished (similarly to the main alignment types of active ditransitive constructions) (Malchukov et. al 2010, 27–28):

1. Indirective passivization: T passivizes, but R does not;
2. Secundative passivization: R passivizes, but T does not;
3. Neutral alignment: either R or T can passivize.

According to the expectations, the alignment of passivization often follows the general alignment of encoding. So if a language uses secundative constructions, most probably it will use a secundative alignment in passivization as well, that is, it will have a preference for R-passivization. In some languages which have alternating ditransitive constructions, only the P-like argument (i.e. the syntactic object) can be passivized from each alignment types. For example, in Northern Khanty R-passivization seems to be possible only from the secundative construction, where R is the syntactic object:

(25) *Pētrāj-na   χōp-na   mōjl-ə-s-a[11]*
      Peter-LOC   boat-LOC   give-EP-PST-PASS.3SG
      'He was given a boat by Peter.'

The fact that only the P-like argument can be passivized from the alternating patterns also shows that there is a connection between the general alignment of encoding and the alignment of passivization (Malchukov et. al. 2010, 28).

However, the alignment of passivization does not follow necessarily the encoding: there are languages with neutral encoding which use a secundative alignment in passivization (it is very common, cf. English examples (26a) and (26b)), and also languages with indirective encoding and a neutral alignment in passivization. One combination is unattested: it seems that there is no

---

[11] In Khanty both the Agent of the passive sentence and the Theme of the SOC construction are indicated with the Locative-Instrumental case.

language which has secundative encoding and strictly indirective passivization. (Malchukov et. al. 2010, 29–30)

(26)   a. *The children were given sweets.*
       b. *? The sweets were given children.*

(Malchukov et. al. 2010, 29)

R-passivization is generally more frequent than T-passivization. The reason for this can be found in the function of passivization, namely the promotion of the topical object. Since in a ditransitive construction R tends to be more topical than T, it is understandable that R-passivization is generally favoured more than T-passivization. (Malchukov et. al. 2010, 30)

In Mansi ditransitive constructions, the ratio of active and passive sentences is far more balanced than in monotransitive ones. Although passive is used in Mansi very frequently, the ratio of passive constructions was under 10% in one of Skribnik's surveys containing 2000 transitive clauses (Skribnik 2001). In our corpus, this ratio was considerably bigger: 39% (see Table 2). It could follow logically from the fact that in a three-argument construction there are two arguments which can rival for the subject position. From a syntactic point of view, in ditransitive constructions the number of the arguments which can be passivized is bigger, while from a pragmatic point of view, there are two arguments which can occupy the topic position. Our corpus still shows that in Mansi mainly the secundative construction is passivized (R-passivization) although there are examples for the passivization of the indirective alignment as well (T-passivization). Preferring R-passivization over T-passivization corresponds with the cross-linguistic evidence (see above).

As it is attested typologically, also in Mansi only the P-like argument can be passivized from both alignment types. Thus passivization of the Mansi secundative construction results in R-passivization (27) and passivization of the indirective construction leads to T-passivization (28).

i)   R-passivization from a secundative construction:

(27)   *Rajonə-t-t*   *ōl-ne*   *pūlńica-t*   *jomas*
       district-PL-LOC   be-PTCP.PRS   hospital-PL good
       tērp-il   *tāstu-wē-s-ət.*
       medicine-INSTR   prepare-PASS-PST-3PL
       'Good medicine was prepared for the hospitals in the districts.'

(LS 2012/22, 4)

ii)   T-passivization from an indirective construction:

(28)   *jārm-ən*   *ta-ke*   *maj-we-s-əm*
       poverty-LAT   that-PTCL   give-PASS-PST-1SG
       'It is poverty that I was given to.' ['It is poverty that I was made to experience.']

(VNGy IV, 330)

Concerning passivization, we also have several examples in our corpus which do not correspond to the expectations based on the information structure (i.e. the most topical element is passivized since the most important function of passivization in Mansi is keeping an important, topic-like element in the subject position c.f. Kulonen 1989, 41, 288). The R-

passivization occurs also quite frequently only because the agent of the verb is indefinite or general and in many of these cases R can not be considered topical.[12]

(29) *Ań   sāw   ńāwram-əŋ   koltāγl-ət   kol   ūntt-ən*
      now   many   child-ADJ   family-PL   house   build-PTCP.PRS
      māγəs   mā-lomt-əl   mi-we-t.
      for   land-piece-INSTR give-PASS.PRS-3PL
      'Now a plot of land is given to the families having many children in order to build a house.'

                                                                                                 (LS 2015/24, 2)

T-passivization appears far more rarely than R-passivization, due to the reason mentioned above: R generally tends to be more topical than T. The motives behind the use of T-passivization seem to be a lot less clear than those guiding the use of R-passivization. There are also such sentences in the corpus, where T-passivization seems to have an emphasizing function, it appears that passivization is used in order to put an extra emphasis on the Theme.[13] See for example (30):

(30) *Sverdlovski oblast'-it   mansi   mir-n   nemater ńotmil*
      Sverdlovsk region-PL   Mansi   people-LAT   nothing help
      at   majl-awe.
      NEG   give-PASS.PRS.3SG
      'For the Mansi people of the Sverdlovsk region is given no help at all.'

                                                                                                 (Dinislamova 2007, 8)

5   Conclusions

5.1   Statistical findings

The following tables summarize our findings. As it can be seen from Table 2, passive constructions are fairly frequent also in ditransitive patterns.

Active	122 (61%)
Passive	78 (39%)
Total	*200 (100%)*

Table 2. *Active and passive ditransitive constructions in Northern Mansi*

Table 3 shows that indirective constructions seem to dominate. We assume, however, that this data should be treated with caution since our further investigation based on a larger database does not show the dominancy of IOCs. It is worth mentioning, that a similar statistics on

---

[12]   Although using a passive construction in case of a generic, unknown or unimportant agent is quite common cross-linguistically, this function of the passive is not typical of Mansi (nor Khanty). One of the unusual properties of Mansi passive is that "Agent demotion or supression is rather a peripheral function – the Agent is often present in the sentence." (Skribnik 2001, 224) In modern Mansi texts (especially in the newspaper), however, it seems that passive has obtained this cross-linguistically common function, too.

[13]   While our selected corpus of 200 ditransitive clauses used for this study contains only 6 examples of T-passivization (cf. Table 6), it is important to note that in our whole, extended corpus there are several other clauses with T-passivization. Thus the statements about the use and the function of T-passivization are based on this extended corpus.

Eastern Mansi ditransitivity shows the dominancy of the secundative constructions (Virtanen 2011).

Indirective (IOC)	76 (62,3%)
Secundative (SOC)	46 (37,7%)
Total	*122 (100%)*

Table 3. *Active ditransitive constructions in Northern Mansi*

Table 4 indicates that in the majority of indirective constructions the object-verb agreement is lacking.

IOC + subjective conjugation	51 (67,1%)
IOC + objective conjugation	25 (32,9%)
Total	*76 (100%)*

Table 4. *Active indirective ditransitive constructions (IOC) in Northern Mansi*

In SOCs the object-verb agreement is decisive what is in correlation with the pragmatic role of this pattern (Table 5).

SOC + subjective conjugation	3 (6,5%)
SOC + objective conjugation	43 (93,5%)
Total	*46 (100%)*

Table 5. *Active secundative ditransitive constructions (SOC) in Northern Mansi*

Similarly the dominance of R-passivization is in accordance with its pragmatic role (Table 6).

R-passivization (passive SOC)	72 (92,3%)
T-passivization (passive IOC)	6 (7,7%)
Total	*78 (100%)*

Table 6. *Passive ditransitive constructions in Northern Mansi*

5.2 Summary

The analysis of ditransitive constructions in a typological frame showed that:

In Mansi ditransitive constructions there is an alternation between indirective and secundative constructions which is a common phenomenon cross-linguistically.

Alternation is controlled primarily by topicality, this is a general phenomenon also cross-linguistically.

Mansi ditransitives prefer R-passivization, this kind of passivization is also the most frequent one in languages.

Our findings thus show that Mansi ditransitive constructions are not unique typologically – as opposed to the previous statements (cf. **1**). Although general tendencies prevail also in Mansi, there are contradictory examples the use of which we intended to explain in our paper (cf. examples 20–21 and 29–30), some of them, however, still remains unexplained (cf. example 17).

Abbreviations

A	agent of a (di)transitive clause
ACC	accusative
ADJ	adjective marker
ART	article
CO	coaffix
DAT	dative
DIM	diminutive
DOC	double object construction
DST	destinative
DU	dual
EMPH	emphatic element
EP	epenthetic element
INSTR	instrumental
IOC	indirect object construction
LAT	lative
LOC	locative
NAR	narrative
O	objective conjugation (in Hungarian)
PASS	passive
PL	plural
PRS	present
PST	past
PTCL	particle
PTCP	participle
R	recipient
SG	singular
SOC	secundary object construction
T	theme
V	verb

References

Beck & Johnson. 2004. Double Objects Again. *Linguistic Inquiry*. Winter (35): 97–123.

Bíró, Bernadett. 2013. Ditranzitív igékből képzett cselekvésnevek a manysiban. [Mansi nomina actionis derived from ditransitive verbs.] *Nyelvtudományi Közlemények* 109. 137–150.

Bíró, Bernadett. 2015. Az északi és déli manysi ditranzitív szerkezetek szórendje. [Word order of ditransitive constructions in Northern and Southern Mansi dialects.] In Katalin É. Kiss, Attila Hegedűs, Lilla Pintér (eds.), *Nyelvelmélet és dialektológia 3*, 43–58. Piliscsaba: PPKE.

Bíró, Bernadett & Sipőcz, Katalin. 2009. 14. Language shift among the Mansi. In James N. Stanford & Dennis R. Preston (eds.), *Variation in Indigenous Minority Languages,* 321–346. Amsterdam-Philadelphia: John Benjamins Publishing Company.

Bresnan, Joan & Cueni, Anna & Nikitina, Tatiana & Baayen, R. Harald. 2007. Predicting the dative alternation. In G. Boume, & I. Kraemer & J. Zwarts (eds.), *Cognitive Foundations of Interpretation,* 69–94. Amsterdam: Royal Netherlands Academy of Arts and Sciences.

Dalrymple, Mary & Nikolaeva, Irina. 2014. *Objects and Information Structure*. Cambridge: University Press.

Dolovai, Dorottya. 2003. Tárgyas mondatszerkezetek a vogulban. [Transitive constructions in Vogul.] In Beatrix Oszkó-Mária Sipos (eds.), *Budapesti Uráli Műhely III. Uráli Tárgyaló*, 63–76. Budapest: MTA Nyelvtudományi Intézet.

Dinislamova, Svetlana. 2007. *Lavim l'oyxanuw.* [The way destined for us.] Khanty-Mansiysk: Poligrafist.

Givón, Talmy. (1984/1990). *Syntax: a functional-typological introduction* (Vol. I, 1984; Vol. II, 1990). Amsterdam & Philadelphia: John Benjamins.

Haspelmath, Martin. 2005. Argument Marking in Ditransitive Alignment Types. *Linguisti Discovery*. 3/1. (https://journals.dartmouth.edu/cgi-in/WebObjects/Journals.woa/1/xmlpage/1/article/280?htmlOnce=yes)

Honti, László. 1969. A tárgy jelölése a vogul nyelv tavdai nyelvjárásában. [Object marking in the Tawda dialect of the Vogul language.] *Nyelvtudományi Közlemények* 71. 113–121.

Honti, László. 1988. Die ob-ugrischen Sprachen. In Denis Sinor (ed.), *The Uralic Languages. Description, History, and Foreign Influences,* 172–196. Leiden – New-York: E. J. Brill.

Honti, László. 1999. Ugor alapnyelv: téves vagy reális hipotézis? In Marianne Bakró-Nagy & Zoltán Molnár & Zsuzsa Salánki & Mária Sipos (eds.), *Budapest Uráli Műhely I. Ugor Műhely*, 19–42. Budapest: MTA Nyelvtudományi Intézet.

Kálmán, Béla. 1976. *Wogulische Texte mit einem Glossar*. Budapest: Akadémiai Kiadó.

Keresztes, László. 1998. Mansi. In Daniel Abondolo (ed.), *The Uralic Languages*, 387–427. London: Routledge.

É. Kiss, Katalin. 2010. A magyar tárgyas és alanyi igeragozás kialakulásának szintaktikai hátteréről. [The syntactic background of the emergence of the Hungarian objective and subjective conjugations.] *Nyelvtudományi Közlemények* 107. 131–146.

É. Kiss, Katalin. 2013. The Inverse Agreement Constraint in Uralic Languages. *Finno-Ugric Languages and Linguistics* Vol. 2. No. 1. 2–21.

Kulonen, Ulla-Maija. 1999. Object Marking in the Ugric Languages. In Marianne Bakró-Nagy & Zoltán Molnár & Zsuzsa Salánki & Mária Sipos (eds.), *Budapest Uráli Műhely I. Ugor Műhely,* 63–72. Budapest: MTA Nyelvtudományi Intézet.

Kulonen, Ulla-Maija. 2007. *Itämansin kielioppi ja tekstejä*. [Eastern Mansi grammar and texts.] Hilfsmittel für das Studium der finnisch-ugrischen Sprachen XV. Helsinki: Suomalais-Ugrilainen Seura.

Lavotha, Ödön.1953. A tárgy jelölése a manysiban. [Object marking in Mansi.] *Nyelvtudományi Közlemények* 54. 200–218.

LS: *Luima Seripos*. Mansi language newspaper. Hanti-Manszijszk. http://www.khanty-yasang.ru/

Malchukov, Andrej & Haspelmath, Martin & Comrie, Bernard. 2010. Ditransitive constructions: a typological overview. In Andrej Malchukov & Martin Haspelmath & Bernard Comrie (eds.), *Studies in ditransitive constructions: A comparative Handbook*, 1–65. Berlin: Mouton de Gruyter.

Margetts, Anna & Austin, Peter. 2007. Three-participant events in the languages of the world: towards a cross-linguistic typology. *Linguistics* 45.3. 393–451.

Nikolaeva, Irina. 1999. *Ostyak*. Munich: LINCOM Europa.

Nikolaeva, Irina. 2001. Secondary Topic as a Relation in Information Structure. *Linguistics* 39.1: 1–49

Rappaport Hovav, Malka & Levin, Beth. 2008. The English Dative Alternation: The Case for Verb Sensitivity. *Journal of Linguistics* 44. 129–167.

Rombandeeva, Evdokija Ivanovna. 1979. *Sintaksis mansijskogo (vogul'skogo) jazyka*. [The Syntax of the Mansi (Vogul) Language.] Moskva: Nauka.

Sipőcz, Katalin. 2011. Az obi-ugor ditranzitív szerkezetek történeti háttere. [The historical background of the Ob-Ugric ditransitive constructions.] In Marianne Bakró-Nagy & Tamás Forgács (eds.), *A nyelvtörténeti kutatások újabb eredményei VI*, 221–229. Szeged: Szegedi Tudományegyetem Magyar Nyelvészeti Tanszék.

Sipőcz, Katalin. 2013. Ditranzitív igék a manysiban. [Ditransitive verbs in Mansi.] *Nyelvtudományi Közlemények* 109. 123–136.

Sipőcz, Katalin. 2015a. Tárgyjelölés a manysi ditranzitív szerkezetekben. [Object marking in Mansi ditransitive constructions.] In Katalin É. Kiss & Attila Hegedűs & Lilla Pintér (eds.), *Nyelvelmélet és dialektológia 3,* 234–246. Piliscsaba: PPKE.

Sipőcz, Katalin. 2015b. Ditransitivity in the Ob-Ugric languages. In Harri Mantila & Kaisa Leinonen & Sisko Brunni & Jari Sivonen & Santeri Palviainen (eds.), *Congressus Duodecimus Internationalis Fenno-Ugristarum. Plenary Papers*, 133–159.Oulu: University of Oulu.

Sipőcz, Katalin. 2016. A manysi ditranzitív igék szintaxisa szemantikai megközelítésből. [The syntax of ditarnsitive verbs from semantic approach.] *Jelentés és Nyelvhasználat* (3). 49–61.

Skribnik, Elena. 2001. Pragmatic Structuring in Northern Mansi. In Tõnu Seilenthal & Anu Nurk & T. Palo (eds.), *Congressus Nonus Internationalis Fenno-ugristarum 2. 7.-13.8.2000 Tartu. Pars VI. Dissertationes sectionum: Linguistica III.*, 222–239. Tartu: Eesti Fennougristide Komitee.

Szilágyi, Norbert. 2014. *A tavdai manysi tárgyjelölés morfoszintaktikai alapú vizsgálata digitalizál korpusz alapján.* [Morphosyntactic investigation of object marking in Tawda Mansi based on digitalised corpus.] Szeged MA: Szegedi Tudományegyetem.

Virtanen, Susanna. 2011. Ditransitiivien vaihtelua mansin itämurteissa. [Ditransitive alternation in Eastern Mansi.] *Folia Uralica Debreceniensia* 18. 165–176.

Virtanen, Susanna. 2013. Contextual function of noun marking in the direct object-marking system of Eastern Mansi. In Márta Csepregi & Kata Kubínyi &, Jari Sivonen (eds.), *Grammatika és kontextus: új szempontok*

az uráli nyelvek kutatásában III. Uralisztikai Tanulmányok 20, 304–317. Budapest: Eötvös Loránd Tudományegyetem

VNGY: Munkácsi, Bernát. 1892–1896. *Vogul Népköltési Gyűjtemény I–IV*. [Vogul Folklore Collection.] Budapest: MTA.

Wagner-Nagy, Beáta & Szeverényi, Sándor. 2013. On the argument structure of the 'Give' verbs in Nganasan and Selkup. *Tomskyj zhurnal lingvisticheskih i antropologicheskih issledovanij*. 2013.1. 27–36.

Bernadett Bíró
University of Szeged
birobernadett9@gmail.com

Katalin Sipőcz
University of Szeged
sipoczk@gmail.com

Interplay between Case, Animacy and Number:
Interpretations of Grammatical Role in Estonian[*]

Merilin Miljan, Elsi Kaiser, Virve-Anneli Vihman

This paper reports on a psycholinguistic experiment investigating information used in incremental interpretation to assign grammatical roles to case-marked nouns. The three core grammatical cases in Estonian – nominative, partitive and genitive – are syntactically ambiguous, in that they may function in various grammatical roles depending on the context. Our study probed what grammatical functions Estonian speakers assign to nouns marked in these cases prior to any knowledge of the syntactic structure of the clause. We investigated the effect of animacy and number, in combination with case-marking, on grammatical role assignment of clause-initial nouns. Results show that these cases are underspecified, and that the interpretation of nouns, when presented without prior context, involves semantic cues like animacy and number, which interact with morphological case to guide the assignment of grammatical functions.

Keywords: *grammatical relations, incremental interpretation, morphological case, Estonian*

1   Introduction

Grammatical role assignment, or building syntactic relationships between words, forms a fundamental part of language comprehension. Building syntactic structure from the words we read or hear takes place *incrementally* (e.g. Altmann & Kamide 1999, Kaiser & Trueswell 2004, Tanenhaus et al. 1995 *i.a.*). That is, a comprehender starts building an interpretation incrementally, before the end of the sentence, at a point when the linguistic information is still potentially syntactically ambiguous. Prior work shows that a variety of constraints guides the extent to which the language comprehension system considers or 'activates' alternative structures.

The incremental nature of language processing raises questions for the role that morphological case-marking plays in the process of grammatical role assignment. Morphological case-marking can encode relations between linguistic elements, such as signaling whether a noun is the subject or object of a (potentially as yet unseen) verb. Indeed, morphological case has often been taken to be a reliable, one-to-one means for indicating the grammatical relation between a case-marked argument and the verb (see, e.g. de Swart 2005, Lamers 2005 or most traditional grammars). This seemingly straightforward interpretation of case, however, depends on a linguistic context, which necessarily includes a verb and possibly other arguments of the verb. Even in a context where the verb is known, a particular case does not always unambiguously signal a specific grammatical relation. This is illustrated in the Estonian examples in (1) and (2). Example (1) shows that when more than one genitive-marked noun occurs in sequence, the modifier function is indistinguishable from the object function, unless we take the discourse context into account; in this case, the question is whether the noun 'cone' in (1)

---

should be interpreted as a possessor, hence a modifier of the noun 'hiding place', or the object of the verb 'to bring' (see Roosmaa et al., 2003, who discuss the ambiguity of the genitive in Estonian).

(1) *Orav   viis   selle   käbi   peidukohta.*[1]
    squirrel.NOM.SG[2] took this.GEN   cone.GEN.SG hiding.place.INE.SG
    a.   'The squirrel took this pinecone to a hiding place.'
    b.   'The squirrel took this to the pinecone's hiding place.'

Another example of ambiguity is in (2), where two arguments occur in nominative case form. As Estonian word order is generally determined by information structure, with topics preceding the verb and foci occurring later, the pre-verbal nominative argument in (2) may turn out to be either subject or object. In other words, without further context, the initial noun may be interpreted as the subject or the object argument.

(2) *Eestlased   vahetasid   välja läti   rahuvalvajad.*
    Estonian.NOM.PL   replaced   out   Latvian.GEN.SG   peacekeeper.NOM.PL
    a.   'The Estonians replaced the Latvian peacekeepers.'
    b.   'The Latvian peacekeepers replaced the Estonians.'

So far, we have been considering whole sentences, where the verb information is assumed to be available during the interpretation process. However, recall that processing is known to be highly incremental: Comprehenders start to build an interpretation early on, as soon as they process the first words, before they have encountered the full sentence. Since Estonian sentences tend to be verb-medial, this means that comprehenders often encounter at least one nominal element before the verb. This brings up the fundamental question of how case morphology guides the relational interpretation of a noun when verb-based information is *not yet present*.

In related work, Bornkessel-Schlesewsky & Schlesewsky also note that interpretational effects driven by nominal elements can be "observable prior to the verb, attest[ing] to the verb-independence of role prototypicality assessment, that is, to an abstract, verb-independent notion of A [agent-like] and P [patient-like] roles" (Bornkessel-Schlesewsky & Schlesewsky 2009, 33).

The question of how case morphology guides noun interpretation is relevant for any model of processing, and especially for languages which allow word order variation (i.e., comprehension cannot rely on word order to identify grammatical role). Currently, the role of morphology in language processing is not yet fully understood, and the present paper aims to contribute to these ongoing explorations.

The experiment reported in this paper investigates the interpretation of morphological case by native Estonian speakers, in order to probe the factors relevant to determining the grammatical role of a constituent and the structure of a sentence. Estonian has several characteristics which make it a particularly fruitful language for this

---

[1]   Example from http://kodu.ut.ee/~kaili/Loengud/Mudelid08/lnotes1.pdf

[2]   The abbreviations for glosses used in this paper are as follows: NOM = nominative, GEN = genitive, INE = inessive, PAR = partitive, ALL = allative, ELA = elative, ILL = illative, ADE = adessive, SG = singular, PL = plural, PST = past tense, PRS = present tense, 3 = third person, 2 = second person, 1 = first person, IMP = imperative, IMPERS = impersonal, NEG = negation marker, PCPL =participle, PRTCL = particle.

investigation. First, word order is variable, which means that linear order does not determine the grammatical role of a noun. For example, one can encounter either subjects or objects early in sentence interpretation. Second (and relatedly), case-marking is the most reliable cue to grammatical roles in Estonian, and thus one might expect comprehenders to be especially attuned to it. Third, although case markers are reliable cues, they do not map uniquely to grammatical roles: ambiguity exists both in the morphological case system and in the mapping between morphological case and syntactic role. (This is discussed in more detail in Section 3). Hence, comprehenders have experience in predicting and adjusting their expectations regarding the function of morphologically case-marked nouns.

How informative are morphological case-markers on their own; or, conversely, how ambiguous are case-markers in the absence of a verb which assigns semantic roles to its arguments? What factors become relevant for interpreting case before the verb is available, when core cases are syntactically ambiguous? What properties of the nominal argument help determine its grammatical role assignment and interpretation? It is well known that preverbal argument interpretation relies on interacting information such as animacy and definiteness (see, e.g. Bornkessel & Schlesewsky 2006, Lamers & de Swart 2012). Yet, determining exactly what type of information interacts with what, and under what conditions, is still a major challenge for both linguistic theory and psycholinguistic explanations.

1.1   Aims of this work

This work investigates whether and to what extent morphological information in case-marked nouns in Estonian affects their interpretation in terms of grammatical relations: How well will a case-marker predict the grammatical role of the noun it marks, independent of any verb? Estonian has three basic grammatical cases (nominative, partitive, genitive[3]), each of which can be used in at least two different grammatical roles (see Section 3); hence, without any linguistic context, all three are syntactically ambiguous. In addition, Estonian has relatively free word order.

We investigated how morphological case-marking, in combination with animacy (animate/inanimate) and number (singular/plural), constrains argument interpretation, in order to see how native speakers' interpretations of nouns and their grammatical roles are influenced by these three factors. As we discuss below, we used an offline production experiment which, in certain key respects, imitates the incremental nature of online comprehension. It allows us to probe how people interpret case-marked nouns, in the absence of any verb information, and hence to ask what information is read off the case-marking itself.

The results can shed light on the kinds of (potential) syntactic relations that comprehenders activate when they see a case-marked noun before any verb-related information is available, e.g. in sentence-initial position. In Estonian, word order is driven by pragmatic constraints, and, although prototypical transitive sentences have SVO word order, word order can and does vary. Hence, case-marking is the most reliable cue for disentangling grammatical relations.

Yet case-marking is not a fully deterministic indication of grammatical role either. For example, in situations involving case homonymy, a noun's case is morphologically

---

[3]   The genitive has also been referred to in the literature as (syntactic) accusative, morphologically realised as genitive in singular and nominative in plural.

ambiguous. Furthermore, even when the case-marking is morphologically unambiguous, a noun on its own in any of the three grammatical cases is syntactically ambiguous (Section 3). For example, genitive case may signal a possessive relation or a direct object, both of which are possible in sentence-initial position. Partitive case may be used to mark either a (partially affected) direct object or a subject in certain constructions. Nominative case is used for both canonical subjects as well as certain affected direct objects, e.g. plural direct objects in declarative sentences and any affected object in imperatives.

In light of the incremental nature of language comprehension, we pose the overarching question of what happens when people encounter a sentence-initial case-marked noun whose syntactic role is ambiguous. From a *structural perspective*, one prediction is that comprehenders will initially prefer interpretations compatible with the minimal amount of structure (e.g. extended Argument Dependency Model in Bornkessel-Schlesewsky & Schlesewsky 2009a,b). For example, nominative-marked nouns will tend to be interpreted as subjects of intransitives, thereby assuming the minimal amount of structure.

From a *frequency-based perspective,* a likely prediction is that nouns will tend to be interpreted in accordance with the most frequent patterns in the language (e.g. Bybee & Hopper 2001, Gries & Divjak 2012, MacDonald 2013, see also Hale 2001 and Levy 2008 for related work). For example, based on what is known about Estonian word order and frequency patterns, it is reasonable to assume that sentence-initial nominative nouns will be interpreted as subjects and sentence-initial genitive nouns as possessors. A frequency-based account is not mutually incompatible with a minimal structure based account, as it could well be that frequent structures also tend to involve minimal structure.

Crucially, our central aim in this paper is not to test or disentangle minimal-structure-based or frequency-based accounts, though we do consider both factors to be important and discuss them in the results section. Prior work on sentence-processing more generally has already yielded considerable evidence in favor of both of these accounts. Instead, our focus here is on how properties of the case-marked noun influence comprehenders' expectations about the noun's grammatical role. Prior work has found that preverbal argument interpretation also relies on other factors such as *animacy* and *definiteness* (e.g. Bornkessel & Schlesewsky 2006, Lamers & de Swart 2012). We chose to investigate animacy in our experiment because its relation to case-marking is an unexplored area in Estonian. In addition to testing whether the *animacy* of the case-marked noun influences its interpretation, we also investigate whether the *number* of the noun (singular/plural) matters. Number distinctions have not been closely examined in the context of grammatical role interpretation, but they are known to be relevant in the assignment of grammatical case in Estonian (Cann & Miljan 2012, Erelt et al. 1993, Rajandi & Metslang 1979). For example, a basic fact about Estonian case-marking is that *singular* affected objects are genitive, whereas *plural* affected objects take nominative case. (We did not investigate definiteness, as Estonian does not mark definiteness with any distinctive obligatory markers, and also because definiteness is a category which is entangled with other features.)

In sum, the general question underlying this study is: Which factors are relevant for noun interpretation in a situation where the core cases are syntactically ambiguous and comprehenders have not yet encountered verb information? Under a view where case-marking acts as a clear, deterministic marker indicating a noun's grammatical role, the prediction is that number and animacy will have no effect on how a case-marked noun is interpreted. However, based on the ambiguities inherent in Estonian case-marking, this outcome seems highly unlikely. Instead, if case-marking is a cue that provides

constraining information about a noun's possible grammatical role, yet does not, on its own, map directly onto a particular grammatical role, then we may well find factors like animacy and number modulating comprehenders' interpretation of case-marked nouns. Our experiment will shed light on (i) how number and/or animacy guide interpretation of case-marking, (ii) the relative contributions of animacy and number for the interpretation of nominative, genitive and partitive case, and (iii) whether these three core cases differ in the extent to which animacy and number guide their interpretation.

2   Prior work on the interplay between case, animacy and number

In this section, we briefly review existing views from theoretical linguistics and psycholinguistics on case-marking, with a particular focus on the effects of animacy and number. We then turn more specifically to prior work on how animacy and number interact with case-marking, with a special focus on Estonian.

2.1   Animacy and case-marking

Although it is not clear whether a noun's animacy status is an independent factor (an 'ontological' category in Dahl's terms, 2008) or an artefact reducible to other factors (see discussion in Rosenbach 2005, 2008, Dahl 2008), animacy clearly has consequences for linguistic encoding. For example, in terms of grammatical role, subject referents in nominative-accusative languages tend to be animate and objects tend to be inanimate (e.g., Aissen 2003, Comrie 1989 *i.a.*). Furthermore, this pattern leads to certain tendencies in case marking (e.g. Comrie 1989, de Hoop & de Swart 2008). For example, animate nouns in the direct object function are cross-linguistically more likely to be overtly case-marked (e.g. by accusative), because this is not their typical role (e.g. Comrie 1989, 128).

Animacy has also been shown to have an independent effect in *syntactic* construction alternations. For instance, Rosenbach argues – based on English constructions with *'s* and *of* (*the king's box* vs. *the box of chocolate*) – that animacy is an independent factor in guiding grammatical variation, as well as a processing factor contributing to efficiency in parsing (Rosenbach 2005, 639). Other work highlighting the effect of animacy in syntactic alternations comes from research on the English dative alternation (Bresnan et al. 2007).

Psycholinguistic studies provide ample evidence showing that the animacy of nouns plays an important role during real-time processing, for example in resolving structural ambiguity (e.g. Lamers 2005, Tanaka et al. 2005, Prat-Sala & Branigan 2000, Christianson & Ferreira 2005, Traxler et al. 2002, 2005, Mak et al. 2002, 2006). The focus has been on how animacy contributes to grammatical function assignment in locally ambiguous structures and to word order during production and comprehension (see e.g. Branigan et al. 2008 for an overview). Most of these studies, however, investigate languages like English that do not have rich case-marking paradigms, and thus their main focus is on syntactic alternations.

Less research has considered animacy paired with case-marking, but work that has been done shows that the interaction between morphological case and animacy in determining grammatical role assignment need not be absolute even in a single language, but rather may vary according to the case information available in a particular construction (Bornkessel-Schlesewsky & Schlesewsky 2009b). A particularly clear example investigated how German speakers interpret nouns, morphologically marked as

nominative, dative or accusative, that could act as syntactic subjects or objects. On the one hand, the interpretation of nouns in constructions with *nominative and accusative* arguments (where the nominative noun must be semantically interpreted as the actor, and the accusative noun as the undergoer) is determined by the case-marking of the arguments (nominative=subject, accusative=object, regardless of animacy, Grewe et al. 2007). On the other hand, in constructions with *nominative and dative* arguments (either of which could in principle be interpreted as the actor *or* the undergoer), the animacy of the nouns plays a key role: Animate entities are more preferred as subject (actor) arguments than inanimate ones, leading to a favored linearisation of animate-before-inanimate (Schlesewsky & Bornkessel 2004, Grewe et al. 2006). This shows that animacy affects interpretation preferences in situations where case-marking is ambiguous, in this case between the actor and the undergoer argument (Bornkessel-Schlesewsky & Schlesewsky 2009b, 43–44).

In sum, although animacy clearly plays a role in guiding the interpretation of case-marked arguments, relatively little is known about 'cue reliability' of case-markers in morphologically rich languages, and what type of information they interact with when this cue is ambiguous.

2.2   Number and case-marking

The relationship between number, grammatical role assignment, and case-marking has not received much attention in theoretical or psycholinguistic work. Prior psycholinguistic studies on number mostly focus on subject-verb agreement effects, not case-marking (see Lago et al. 2015 for a review). Yet the same factors that interact with grammatical role assignment have also been found to play a role in number agreement: animacy, definiteness, specificity/referentiality, topicality (e.g. Corbett 2000, Bamyaci et al. 2014).

In Estonian, object case-marking crucially involves number: In predicates denoting unbounded events and/or unaffected objects, *both singular and plural objects* are partitive; as well as objects in the scope of negation (3c). However, for transitive, affirmative predicates involving affected (countable) objects and bounded events, objects have heterogeneous case-marking *depending on number and clause type*, as shown in (3a). Singular affected objects in unmarked, active voice constructions take genitive case, whereas nominative case marks plural affected objects as well as affected objects in subjectless constructions, such as imperatives (as in 3b) and impersonals (Erelt et al. 1993). (The count/mass distinction as well as numeral constructions also play an important role in the use of the partitive, but this paper only focuses on count nouns without numerals.)

(3) a. *Poiss   kirjutas   luuletuse / luuletused.*
         boy.NOM.SG   wrote   poem.GEN.SG/poem.NOM.PL
         '(The) boy wrote (a) poem / poems.'

     b. *Loe   luuletus /   luuletused   ette!*
         read.IMP.2SG   poem.NOM.SG/poem.NOM.PL   forward
         'Recite the poem!'

     c. *Poiss   ei   kirjutanud luuletust / luuletusi.*
         boy.NOM.SG   NEG   write.PCPL   poem.PAR.SG/PAR.PL
         '(The) boy did not write (a) poem / poems.'

3   Core cases and word order in Estonian

3.1   The core cases

Estonian has a rich system of morphological case-marking, including differential marking of grammatical relations and the variable use of nominal case-marking. The canonical subject-marking form, nominative, is also used for objects in various constructions (Section 2.2); and some subjects are partitive. Likewise, partitive and genitive – the object cases[4] – are also used to mark other syntactic relations, e.g. adverbial and possessor, respectively. Thus, there is no one-to-one mapping between morphological case and grammatical function. All three core cases can give rise to differing interpretations, and each grammatical role can be instantiated with more than one case.

   This ambiguity is particularly evident with nouns in sentence-initial position. For instance, upon encountering the clause-initial, nominative noun *õpetajad...* 'teachers' (as in 4), the listener does not know whether it will function as the object, as in (4a) or the subject, as in (4b) – because at the point where a listener or reader encounters the sentence-initial noun, the rest of the sentence has not yet been seen/heard. This ambiguity also occurs with nouns in other cases: the partitive noun *külalisi* 'guests' in (5) may turn out to be the subject (5a) or the object (5b), and the genitive noun *sõbra* 'friend' in (6) may turn out to be a possessor (6a) or a direct object (6b). Note that the partitive subject interpretation is primarily available to plural and mass nouns, as in (5).[5]

(4)   *Õpetajad...*
       teacher.NOM.PL
       a.   *...kutsuti   koosolekule.*
             invite.PST.IMPERS   meeting.SG.ALL
             '(They) invited (the) teachers to a meeting.'

       b.   *...läksid   koosolekule.*
             go.PST.3PL   meeting.SG.ALL
             '(The) teachers went to a meeting.'

(5)   *Külalisi...*
       guest.PAR.PL
       a.   *...saabub   igast   ilmakaarest.*
             arrive.PRS.3SG   every.SG.ELA   cardinalpoint.SG.ELA
             'Guests are arriving from every corner of the world.'

       b.   *...on kutsutud   saja   ringis.*
             is   invite.PST.PCPL   hundred.GEN   about
             'About one hundred guests have been invited.'

---

[4]   Note that the functions that are typically marked by accusative in other languages are marked by genitive in Estonian; that is, there is no morphologically distinct accusative case in Estonian.

[5]   This is not a grammatical constraint, as pragmatically felicitous contexts can be found for count nouns as singular partitive subjects (for some discussion of partitive nouns, see, e.g., Cann & Miljan 2012; Huumo & Lindström 2014; Rajandi & Metslang 1979).

(6)   *Sõbra…*
      friend.GEN.SG

    a.   *…ema*   *helistas.*
         mother.NOM.SG   call.PST.3SG
         'A friend's mother called.'

    b.   *…kutsus*   *ta*   *ikka*   *kaasa.*
         invite.PST.3SG   3SG.NOM   EMPHATIC.PRTCL along
         'S/he still invited a friend along.'

Prior work on Estonian argument-encoding has mainly focused on verb semantics, aspect and the (in)determinate quantity of the noun (e.g. Erelt et al. 1993, Metslang 2013, Tamm 2004). According to the reference grammar of Estonian (Erelt et al. 1993), the verb determines the case of its arguments and object cases depend on the lexical meaning of the verb (see also e.g. Tamm 2004). Metslang (2013, 26) highlights that the most relevant semantic factors for argument encoding in Estonian are the determinate and indeterminate quantity of the nominal referent.

    Animacy has not been investigated as a relevant factor in grammatical role assignment in Estonian; neither has animacy been invoked in descriptions of case assignment rules in Estonian.[6] As the core case markers themselves are 'unreliable cues' for grammatical role assignment, we expect other types of information to be used, potentially including animacy.

3.2   Word order

Estonian has flexible word order. The basic word order is SVX (Koptjevskaja-Tamm & Wälchli 2001), but Tael (1988) found that only 25% of sentences in her written corpus had SVX order, and 24% had XVS order (cf. footnote 8 in Section 5.1). She ascribes this to an interaction of various factors affecting word order, including sentence type, properties of the noun phrase, predicate semantics, information structure, and a tendency toward verb-second order. From an information-structural perspective, the topic (or given information) normally precedes the comment (or new information). Indeed, all else being equal, placing new information somewhere other than the final position yields an infelicitous word order (Erelt et al. 2007, 524). However, various focusing particles (e.g. *ka* 'also, too'; the emphatic particle *ju*; *isegi* 'even'; *hoopis* 'on the contrary'), enable a speaker to indicate new information elsewhere than the final position.

4   Experiment

We conducted a sentence-completion study to investigate whether and how the case-marking of sentence-initial nouns – in nominative, partitive and genitive case – interacts with number and animacy in guiding what grammatical relations Estonian speakers assign to these nouns. On each trial, participants saw a sentence-initial noun, which they used as

---

    [6] In related work on illative and allative cases, two semantic case markers that we do not investigate, Kittilä (2005) shows that in Finnish (closely related to Estonian), the encoding of indirect objects with the goal function depends on animacy: animate goals occur in allative, while the inanimate ones are marked by illative.

the starting point for their sentence, as illustrated in (7). Participants could write any sentence that came to mind. Examples (a-c) show some possible continuations for the prompt word in genitive case in (7). As can be seen here, flexible word order in Estonian means that sentence-initial nouns are not restricted to acting as the subject of the sentence.

(7)   ***Sõbra…***
        'friend.GEN.SG'…

    a. **[*Sõbra*   *ema*]**_{NP}   *tuli*   *külla.*
        friend.GEN.SG   mother.NOM.SG   come.3SG.PST   to-visit
        '(A) friend's mother came to visit.'

    b. **[*Sõbra*]**_{NP}   *kutsus*   *ta*   *pulma.*
        friend.GEN.SG   invite.3SG.PST   3SG.NOM   wedding.SG.ILL
        'S/he invited a friend to the wedding.'

    c. **[*Sõbra*   *juures*]**_{PP}   *vaatasid*   *nad*   *filmi.*
        friend.GEN.SG   at   watch.3PL.PST   3PL.NOM   film.PAR.SG
        'At (their) friend's house, they watched a film.'

In work using sentence-completion methodology, it is widely assumed that participants' production patterns (in this case, what grammatical role to assign to the noun) provide an indication of people's interpretation preferences (e.g. Trueswell et al. 1993, Snedeker & Trueswell 2004 for related work). When faced with a case-marked noun (or any other kind of linguistic element), participants first have to interpret it before they can provide their continuation. Thus, frequent continuations reflect frequent interpretations. Essentially, completion preferences provide a window into interpretation preferences. Sentence-completion methodology has been successfully used within psycholinguistics to tap into a variety of language-processing related questions.

Crucially, we chose to test sentence-initial nouns because this approach allows for the noun to be presented without any information about the verb or the syntactic structure of the clause. This is desirable, because it means that these factors do not limit or guide the interpretation of case. Bearing in mind our aim of obtaining a measure of how case-marking, animacy and number guide speakers' interpretations of the noun, we want to control and minimize the impact of other potential limiting factors.[7]

---

[7] All the cases used in the experiment, when used in sentence-initial position, can in principle signal more than one grammatical function. It is worth noting, however, that our design does constrain people to using the noun in sentence-initial position (see Kaiser, Miljan & Vihman (under review) for a follow-up study which tests nouns that are not constrained to initial position). Although nouns in a variety of grammatical roles can occur in sentence-initial position in Estonian, some grammatical roles – e.g. subjects – are more frequent sentence-initially than others. We discuss this in more depth below.

4.1 Method

4.1.1 Participants
Forty-two adult native speakers of Estonian participated in our sentence-completion study via the Internet.

4.1.2 Materials and design
We manipulated the number (singular, plural) and case-marking (nominative: NOM, partitive: PAR, genitive: GEN) of the initial noun. This yields the six variants shown in Table 1. We also manipulated the animacy of the noun (animate, inanimate). This was done between-items: Each item either used an animate noun (e.g. 'mouse' in Table 1) or an inanimate noun (e.g. 'pencil' in Table 1). (We also tested both count and mass nouns, but discuss only count nouns here.)

NOUN	ANIMATE 'MOUSE'	CASE FORM	INANIMATE 'PENCIL'
SG	*hiir*	NOM.SG	*pliiats*
	hiirt	PAR.SG	*pliiatsit*
	hiire	GEN.SG	*pliiatsi*
PL	*hiired*	NOM.PL	*pliiatsid*
	hiiri	PAR.PL	*pliiatseid*
	hiirte	GEN.PL	*pliiatsite*

Table 1: *Examples of animate noun* hiir *'mouse' and inanimate noun* pliiats *'pencil'*

We used a Latin-square design, common in psycholinguistic research, which ensures that each participant would see each specific noun (e.g. 'mouse') only once, but would see equal numbers of nouns in all six conditions. Furthermore, each participant saw equal numbers of animate and inanimate nouns.

In all, our study included 18 target count nouns, which are the data we present in this paper. We also tested 9 mass nouns and 6 time expressions, but we do not consider that data here. In addition, the study included 32 filler items, which involved a range of different cases and parts of speech (e.g. *kaua* 'long.time', *pargis* 'park.INE', *kiiresti* 'fast', *suuri* 'large.PAR.PL').

4.1.3 Procedure
The sentence-completion task was conducted over the Internet using LimeSurvey (https://www.limesurvey.org/). Participants saw nouns and were asked to write a sentence beginning with the noun provided. Thus, the task is fairly unconstrained: participants can write whatever comes to mind, as long as it starts with the initial word we provided. As mentioned above, this kind of sentence-completion task taps into grammatical-role expectations triggered in participants' minds by information from case-marked nouns, in a context where no information is available about the verb or clause structure. Participants were not told beforehand that our aim was to investigate case or grammatical roles, and their comments regarding the experiment suggest that they remained unaware of our aim even as they completed the experiment.

4.2   Data analysis

We analyzed the grammatical role that participants assigned to the case-marked noun in their continuation sentence. The data was coded by two Estonian speakers, and 28% of the data was fully double-coded to ensure consistency across coders in their use of the coding labels.

A number of coding categories were used to label the grammatical roles assigned to the case-marked nouns when used in sentences. The most frequent grammatical role labels are in Table 2. Examples from our data are in (8)-(17).

CATEGORY LABEL	GRAMMATICAL ROLES
subj	Subject, e.g. (8)
part-subj	Partitive Subject, e.g. (9)
poss-subj	Possessor of the subject, e.g. (10)
obj	Object, e.g. (11)
obj-adv	Object inside a gerund clause, e.g. (12)
obj-pp	Object of an adposition, e.g. (13)
poss-obj	Possessor of an object, e.g. (14)
poss-adv	Possessor of an adverbial, e.g. (15)
compx-obj	Object of a complex structure, e.g. (16)
compx-subj	Subject of a complex structure, e.g. (17)

Table 2: *Summary of the main coding labels used in data analysis*

(8)   *Porgand…*   (carrot-NOM.SG)
      Porgand   on   väga   tervislik.   [subj]
      carrot.NOM   is   very   healthy
      '(A) carrot is very healthy.'

(9)   *Poisse....*   (boy-PAR.PL)
      Poisse   tuli   joostes   juurde.   [part-subj]
      boys.PAR.PL   come.3SG.PST running   to-nearby
      'More boys were running to join in.'

(10) *Rebase…*   (fox-GEN.SG)
      Rebase   saba   on   kohev   ja   ilus.   [poss-subj]
      fox.GEN.SG   tail.NOM.SG is   fluffy.NOM and   pretty.NOM
      '(A) fox's tail is fluffy and beautiful.'

(11) *Jänest…*   (rabbit-PAR.SG)
      Jänest   [ajasid taga] rebane,   hunt   ja   karu.   [obj]
      rabbit.PAR.SG chase.3PL.PST fox.NOM   wolf.NOM and   bear.NOM
      '(A) fox, wolf and bear chased (the) rabbit.'

(12) *Medalit…*   (medal-PAR.SG)
      [*Medalit   saades*] *on alati   hea   tunne.*   [obj-adv]
      medal.PAR.SG   getting   is always   good.NOM   feeling.NOM
      'One always feels good when receiving a medal.'

(13) *Pildi…* (picture-GEN.SG)
Pildi *peal* *oli* *konn.* [obj-pp]
picture.GEN.SG on.ADE be.3SG.PST frog.NOM.SG
'In the picture, there was (a) frog.'

(14) *Sõbra…* (friend-GEN.SG)
Sõbra *nahka* *ei* *maksa* *koorida.* [poss-obj]
friend.GEN.SG skin.PAR.SG NEG pay.PCPL peel
'It is not wise to cheat your friend.' (lit 'it doesn't pay to peel a friend's skin.')

(15) *Kaardi…* (card-GEN.SG)
Kaardi *tagaküljel* *oli* *hinnasilt.* [poss-adv]
card.GEN.SG reverse.SG.ADE be.3SG.PST price.tag.NOM.SG
'On the reverse of the card, there was a price tag.'

(16) *Raamatute...* (books-GEN.PL)
Raamatute *lugemine* *avardab* *mõttemaailma.* [compx-obj]
book.GEN.PL reading.NOM broadens mental-world.PAR.SG
'Reading books broadens the mind.'

(17) Vahu… (foam-GEN.SG)
Vahu *pulbitsemine* *pani* *koka* *tegutsema.* [compx-subj]
foam.GEN bubbling.NOM made chef.GEN.SG act
'(The) bubbling of the foam made the chef act.'

It is worth noting that Estonian has a range of complex nominalization structures, such as the ones exemplified in (16) and (17) above. Thus, the prompt noun could be used as the subject, object or modifier inside a complex nominalized structure. An example of compx-mod is given in (18) below. The grammatical role coded as obj-adv in (12) above differs from compx-obj in that the former occurs inside a gerund clause rather than inside a noun phrase.

(18) compx-mod = modifier in the complex NP
[[**Sajandi** *parima* *lillesordi*] *valimine*] *lõppes…*
century.GEN.SG best.GEN.SG flower.type.GEN.SG choosing.NOM finished
'(The) selection of the best flower of the century was completed…'

In the section below, we report the most common grammatical roles for each of the case-markers. The other, less frequent grammatical roles are grouped together in the 'other' columns in each graph for ease of presentation.

5   Results and discussion

In this section, we present the results separately for each of the three case-markers, in order to highlight the different ways in which they interact with animacy and number. We also investigate the frequency of negation (Section 5.2.1) and the distribution of transitive and intransitive sentences in our data (Section 5.1.1).

To analyze the data statistically, we used mixed-effects logistic regression models (R Core Team, http://www.R-project.org/). Mixed-effects logistic regression is better suited for this kind of categorical data than analyses of variance. We analyzed the sentence continuation patterns for each case separately, and tested for effects of number (singular/plural) and animacy (animate/inanimate). A p-value of 0.05 or less is regarded as significant. Our mixed-effect models used the maximal random effect structure justified by the design and supported by the data.

5.1   Nominative nouns

Figure 1 shows the grammatical roles that participants assigned to nominative nouns when writing their sentence completions. As can be seen in Figure 1, there is a very strong bias to interpret NOM nouns as subject, regardless of animacy and number: roughly 90% of nominative nouns are interpreted as subjects, across both conditions.
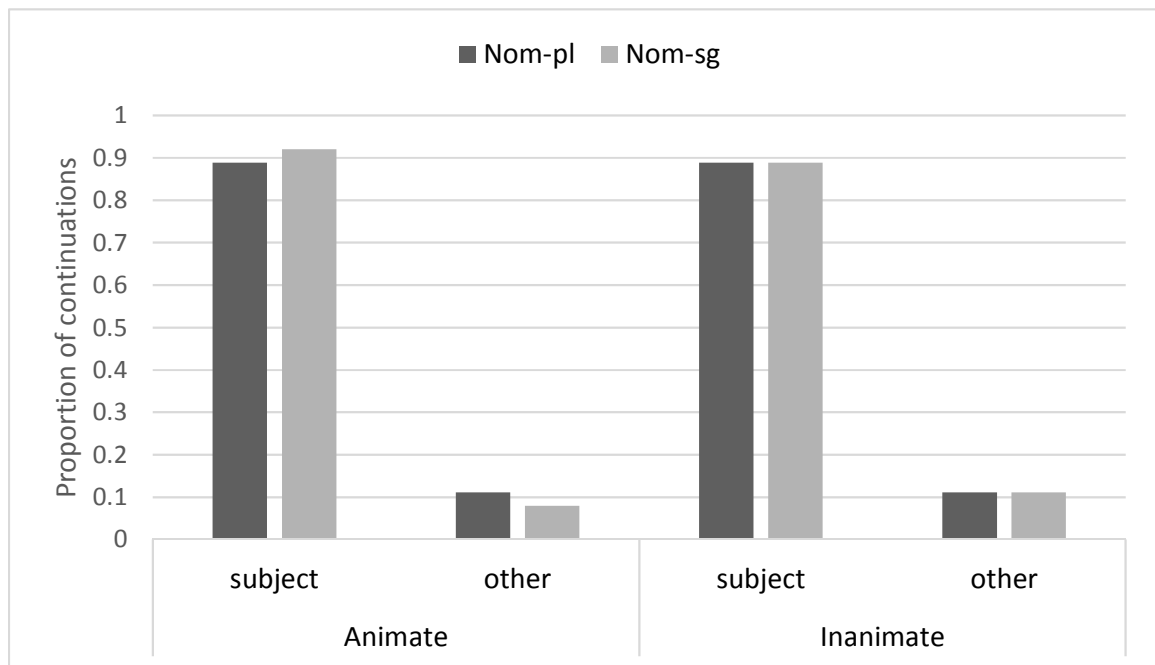
Figure 1: *Grammatical role assigned to nominative nouns*

The strong bias to interpret nominative nouns as subjects is presumably strengthened by the sentence-initial position. Although Estonian has flexible word order, SV(X) is the most frequent core constituent order. While, as noted above, Tael (1988) found only 25% of sentences in a *written* corpus had SVX order, Lindström (2004) reports that in *spoken* Estonian, 41% of all sentences have SV(X) order. Results of our own (ongoing) corpus study indicate that 45% of sentences in the written data and 55% of utterances in the spoken corpus are subject-initial.[8] It is clear that subject-initial order

---

[8]   This is currently based on analysis of 1510 sentences: 752 from a fiction corpus (a subcorpus of 5 million words in the University of Tartu's Balanced Corpus of Written Estonian; online search engine at http://cl.ut.ee/korpused) and 758 from a spoken dialogue corpus (from the University of Tartu's Corpus of Spoken Estonian, maintained by the research group of Spoken Estonian). The difference between our numbers and Tael's may be due in part to differences in the kinds of texts that were included in the corpora. Note that our figures above are for S(X)V(X).

is frequent, though not overwhelmingly so. Importantly, *non-subject-initial* sentences tend to have initial adverbials (32%, according to Tael), rather than other core arguments. Our finding that nominative, sentence-initial nouns are predominantly interpreted as subjects fits well with data based on corpus frequencies.

It is worth noting at this point that a growing body of psycholinguistic research suggests that frequency effects pervade the language processing system in many ways (e.g. Hale 2001, Levy 2008, MacDonald 2013). Thus, our finding that our participants interpreted nominative nouns in a way that matches the frequency patterns of Estonian is not unusual, and not specific to offline tasks. Online methods also show that the human language processing system is highly attuned to the frequency patterns of human languages on a variety of levels (Hale 2001, Levy 2008, see also Wu et al. 2012).

5.1.1 Transitive and intransitive sentences with nominative nouns

The strong subject bias we observe with nouns in nominative allows us to ask a follow-up question regarding potential effects of a minimality bias. More specifically, prior work has argued that people should have a preference for interpretations that are compatible with minimal structure (e.g. the extended Argument Dependency Model in Bornkessel-Schlesewsky & Schlesewsky 2009a,b). For example, nouns in unmarked form, i.e. nominative, are predicted to be interpreted as subjects of intransitives rather than transitives, as intransitives involve less structure.
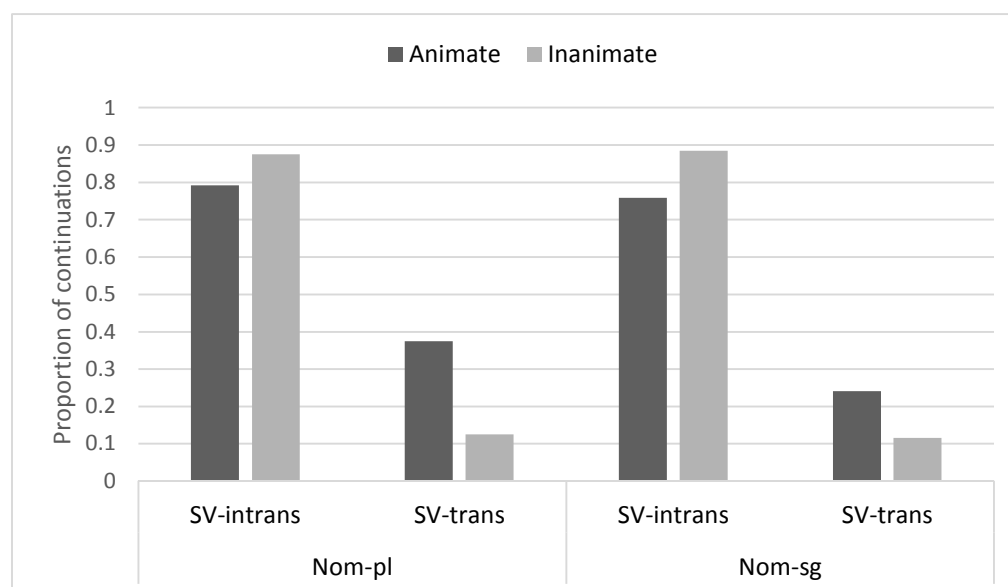

Figure 2: *Proportion of transitive and intransitive continuations with SV and SVO word order elicited by nominative nouns*

We tested this prediction with nominative nouns by comparing the proportions of intransitives with SV word order and transitives with SVO word order. As Figure 2 shows, intransitive continuations are much more frequent than transitive continuations, with both plural and singular, animate and inanimate nouns. This pattern supports the minimality-based idea that there is a preference for minimal structure.

It is important to note, as we show in more detail in Section 6, that this prevalence of intransitives seems to indicate a preference for intransitive over transitive argument structure, and cannot be blamed on 'lazy participants', i.e. it is *not* due to participants

simply writing short sentences in order to finish the task quickly. As we show in Section 6, participants often write rather lengthy, complex continuations.

5.2   Partitive nouns

Figure 3a shows the grammatical roles assigned to partitive nouns in participants' sentence completions. As can be seen in the figure, partitive nouns show considerably more sensitivity to number and animacy than nominative nouns. Animate partitive nouns are most frequently interpreted as objects (singular and plural) and partitive subjects (plural only), and inanimate partitive nouns are most frequently interpreted as objects, partitive subjects or adverbial complements (e.g. see ex. 12 above, obj-adv). Let us assess these three most common continuation types statistically, to determine whether their frequency is significantly influenced by animacy or number.
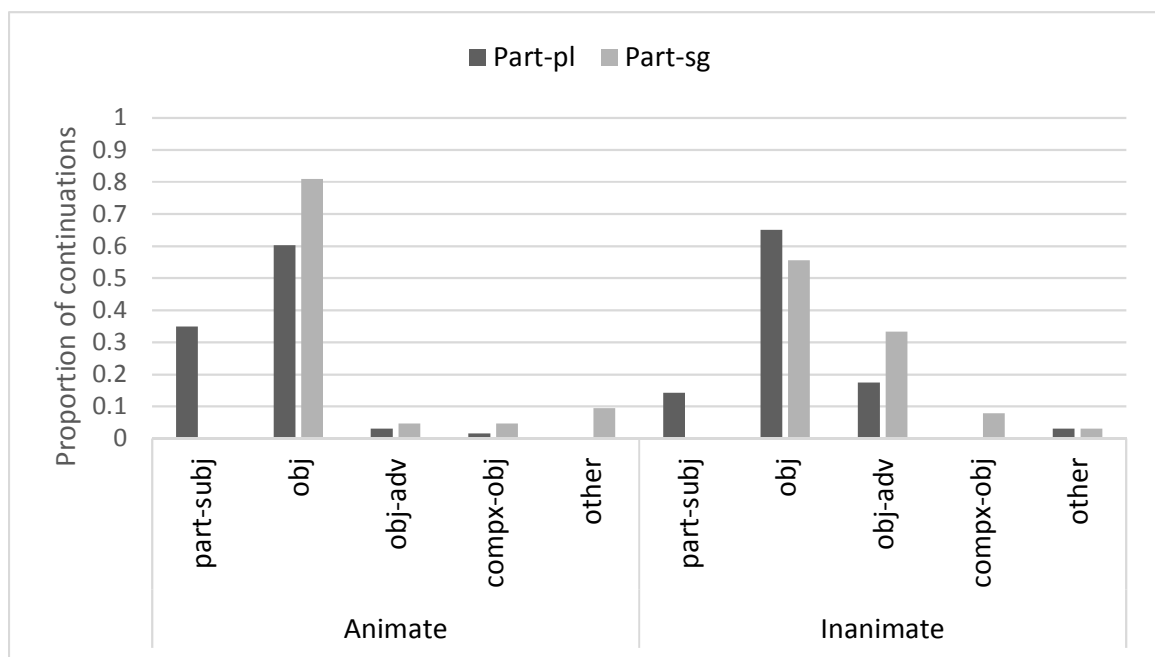
Figure 3a: *Grammatical role assigned to partitive nouns*

When we assess the rate of *object continuations* statistically, we find a main effect of animacy ($\beta = 0.337$, $z = 2.024$, p<.05), no main effect of number ($\beta = 0.23$, $z = 1.137$, p>.25) and a number-by-animacy interaction ($\beta = 0.58$, $z = 3.463$, p<.001). Planned comparisons reveal that *animate nouns* result in significantly more object interpretations when they are singular than when they are plural ($\beta = 0.8102$, $z = 3.138$, p<.002). We suggest that this is best interpreted as a consequence of the availability of *subject* interpretations with plural vs. singular nouns. More specifically, as mentioned in Section 3.1, animate *plural* (count) nouns in the partitive case can easily be interpreted as subjects, but animate *singular* (count) nouns in the partitive case can only be interpreted as subjects in specific contexts.[9] The availability of the subject interpretation with animate partitive

---

[9]   It is important to note that this is not a grammatical constraint, but a pragmatic one. Partitive seems to coerce a mass noun interpretation on nouns which are not the direct object of a verb, and one needs to construct felicitous contexts for singular partitive-marked count nouns to be interpreted as subjects.

plural nouns seems to be why the rate of object interpretations with those nouns is lower than with animate partitive *singular* nouns.

With *inanimate nouns*, the rate of object interpretations is not significantly influenced by the singular/plural distinction (β = -0.269, z = -1.298, p>.19). This is not surprising, for at least two reasons. First, inanimate nouns are crosslinguistically well-known to be less suitable 'subject candidates' than animate nouns – as is also reflected by the lower rate of subject interpretations with inanimate partitive plural nouns, compared to animate partitive plural nouns, in Figure 3a – and consequently an inanimate plural partitive noun is less likely to be interpreted as a subject (and thus less likely to generate the number asymmetry seen with animate nouns). Second, partitive inanimates – especially singular nouns – are also often interpreted as adverbial complements. In essence, on a singular inanimate noun, partitive case is not a strong cue for any one single grammatical role.

When we look more closely at the rate of *adverbial complements (i.e. obj-adv) continuations,* we find that the likelihood of these continuations is not influenced by the number of the noun (β = 0.301, z = 0.854, p>.39) but does show significant effects of animacy (β = -3.188, z = -2.541, p<.02), and no significant number-by-animacy interaction (β = -0.222, z = -0.629, p>.5). As can be seen in Figure 3, adverbial complement (obj-adv) continuations are significantly more likely with inanimate nouns than animate nouns, regardless of number. In our view, this fits well with cross-linguistic observations about objects being prototypically inanimate (all else being equal[10]), and suggests that this link between objects and inanimates extends beyond objects of the matrix verb and also applies to objects inside gerund clauses, even in sentence-initial position.

When we look more closely at the rate of *partitive subject continuations* (in this case, only with plural count nouns), we find that subject continuations are significantly more frequent with animate (plural) nouns than inanimate (plural) nouns (β = 0.564, z = 2.434, p<.02). In other words, *animate* partitive plural nouns are more likely to be interpreted as (existential/presentational) subjects (as in 9 above) than are *inanimate* partitive plural nouns. This finding suggests that the cross-linguistic observations regarding correlations between subjects and animacy also extend to presentational subjects and are not limited to canonical agentive subjects.

In sum, these patterns show that Estonian speakers' expectations regarding the grammatical role of a sentence-initial partitive noun is sensitive to *both* number and animacy. In particular, the overarching object preference is modulated by *number* of the noun if the noun is *animate*: partitive plural nouns are also likely to be interpreted as subjects. We attribute this to a general preference to interpret animate nouns as subjects.

5.2.1 Negation and partitive nouns
Given that objects in the scope of negation in Estonian are always in the partitive case, an obvious question arises regarding the frequency of negative sentences with partitive prompt nouns. In particular, can the high rate of objects that we observed in Figure 3a be attributed to partitive case creating an expectation that the partitive noun is the object in a negated predicate.

---

[10] It is worth noting that although inanimates crosslinguistically tend to be 'non-subjects', there are multiple cues to grammatical role in a rich case-marking language like Estonian (indeed, that is the focus of our paper). Thus, we do not expect inanimates to necessarily show an overwhelming object preference across the board, especially in light of the fact that Estonian also has a variety of other constructions in which partitive inanimates can be realized in non-subject position.

This is worth investigating, because it has implications for the directness of the relation between case and grammatical role expectations. If the object continuations that participants produce with partitive case are mostly in the form of negative sentences, this might suggest that partitive case leads people to expect negation, which in turn leads them to expect a partitive-marked object, rather than partitive case directly triggering an expectation of the noun acting as an object. To look into this, we investigated what subset of the data, shown in Figure 3b, involves negative sentences.
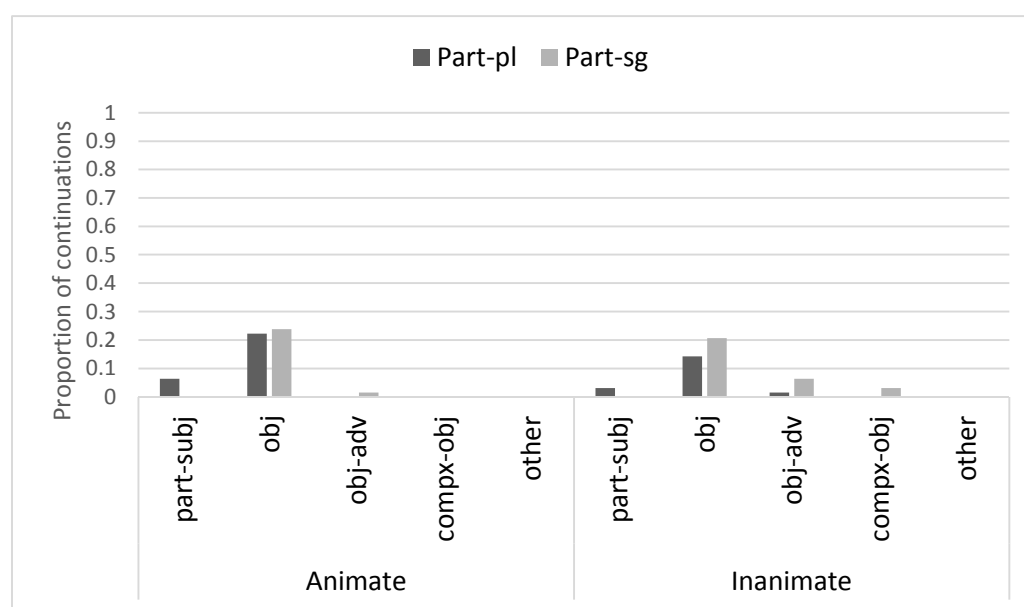

Figure 3b: *The proportion of negative sentences produced by participants as a function of the number, animacy and grammatical role assigned to partitive nouns*

Figure 3b shows the proportion (subset) of negative sentences, relativized to the proportion of continuations for the different grammatical roles. The key point to note is that when we restrict ourselves to the subset of the data that contains just negative sentences, we see that with both animate and inanimate nouns, the proportion of object continuations is *relatively low*.

The fact that negative sentences are not to 'blame' for the patterns we obtained can be seen more specifically if we consider the 60% object continuations with *animate* plural nouns and the 81% object continuations with animate singular nouns (in Figure 3a), these numbers are made up of 22% negative and 38% affirmative sentences (to make up a total of 60% objects with animate plurals), and 24% negative sentences and 57% affirmative sentences (to make up 81% objects with animate singular nouns) respectively. Thus, in both cases, most object continuation sentences are *not negative*.

A similar pattern obtains with *inanimate* nouns: if we consider the 65% and 56% object continuations seen overall, with inanimate plural nouns and inanimate singular nouns respectively, these numbers are made of up 14% negative + 51% non-negative sentences, and 21% negative + 35% non-negative sentences respectively. Again, most of these sentences are *non-negative*. This means that the patterns we obtained are not triggered by the association of partitive case marking with object arguments in the scope of a negated verb.

In sum, the high rate of object proportions that we observe in the dataset of partitive-marked nouns as a whole (shown in Figure 3a) is not an artefact of participants

producing a high rate of negative sentences. Rather, an interaction between animacy, number, and partitive case-marking is observed.

5.3   Genitive nouns

The grammatical roles that participants gave to genitive nouns in their sentence continuations can be seen in Figure 4. (Recall that genitive is one of the core object cases in Estonian, which has no morphologically distinct accusative case.) With genitive nouns, we see that both animacy and number influence the interpretation of the case-marked noun. Animate genitive nouns are likely to be interpreted as possessors of the subject at a higher rate than inanimate nouns. Indeed, as can be seen in Figure 4, none of the animate nouns was assigned the grammatical role of (direct) object. On the other hand, inanimate nouns show higher rates of compx-obj modifier continuations than animate nouns. Compx-obj structures embed the prompt noun as an object inside a complex nominalized structure (as in 16 above). Let us now assess the statistical significance of these patterns.
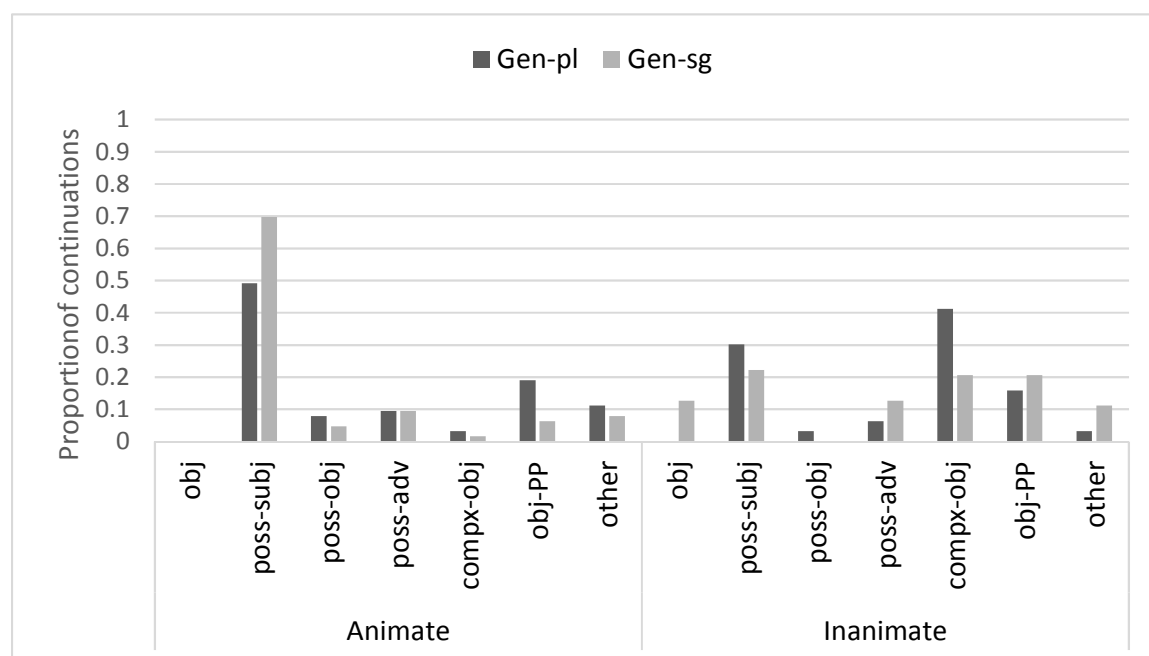

Figure 4: *Grammatical role assigned to genitive nouns*

The proportion of *possessor-of-subject continuations* shows no significant effects of number ($\beta = 0.188$, $z = 0.964$, p>.03), but does reveal a significant effect of animacy ($\beta = 0.989$, $z = 4.396$, p<.0001) and a significant number-by-animacy interaction ($\beta = 0.415$, $z = 2.047$, p<.05). More specifically, there are more possessor-of-subject continuations with animates than inanimates. Furthermore, with *animates,* the rate of possessor-of-subject continuations is marginally higher with singular than plural nouns ($\beta = 0.512$, $z = 1.733$, p<.083), whereas with *inanimates,* number has no significant effect ($\beta = -0.309$, $z = -1.243$, p>.2). In sum, animate singular nouns are the most likely to be interpreted as possessors of the subject of the sentence. This pattern can perhaps be attributed to a dispreference for associating animate nouns with an object role and a preference for associating them with the subject role: although a genitive-marked noun

cannot be interpreted as a subject, it can be interpreted as the possessor of the subject, i.e., as part of the larger subject constituent.

The rate of *compx-obj continuations* shows an effect of animacy (β= -4.634, z = -1.979, p<.05) but no effect of number (β = -0.87, z = -1.386, p>.17) and no interaction (β = -0.71, z =-1.167, p>.2). Thus, compx-obj continuations occur more with inanimates than animates, regardless of number. Given that compx-obj continuations embed the noun as an object inside a complex nominalized structure, this pattern makes sense in light of the cross-linguistically observed association between objects and inanimates. In general, the two most frequent continuation types with inanimates are poss-subj and compx-object.

It is worth noting that only 13% of genitive singular inanimate nouns were interpreted as the direct object of a transitive verb (what some would analyse as accusative), which may seem surprisingly low, considering that genitive marks prototypical highly affected objects. This low rate may be related to word order effects related to genitive in Estonian: whereas sentence-initial objects do occur, they are very rarely genitive case-marked objects.[11] The initial position may have led our participants to embed genitive marked nouns in complex subject constituents. Further study is needed to tease out the effect of word order from the likelihood of assigning an object role to a genitive case-marked noun (see Kaiser, Miljan & Vihman (under review)).

6   General discussion and conclusions

In this paper, we report on a sentence-continuation experiment that investigated how native Estonian speakers interpret syntactically ambiguous, case-marked nouns. We used nominative, genitive and partitive case-marked nouns and manipulated animacy (animate/inanimate) and number (singular/plural), in order to investigate whether and how these factors influence the grammatical role speakers assign to the nouns. We used an offline production experiment that imitates the incremental nature of online processing, designed to elicit interpretations of case-marked, verb-independent nouns, in order to determine how much information is read off of the case-marked noun itself.

Broadly speaking, our results show that morphological partitive and genitive case-markers do provide information about the grammatical role interpretation of a noun. This is supported by the fact that the nouns marked by these cases show biases for grammatical roles, distinct from that preferred by the nominative nouns, which are (in effect) unmarked for case. Unmarked case forms tend to be interpreted as subjects, regardless of animacy, as has been previously demonstrated (e.g. Demiral et al. 2008, Bornkessel-Schlesewsky & Schlesewsky 2009a,b).

The interpretation of genitive and partitive case shows significant effects of animacy and number in determining the more precise functions of case-marked nouns. This underscores the syntactic ambiguity of these case-markers and points to an analysis whereby morphological case-markers should be taken as underspecified: the more precise syntactic function of a case-marked noun is determined through interpretation of the case, animacy and number of the noun, even without the context. For example, genitive

---

[11]   In our ongoing analysis of written and spoken Estonian corpus data (see footnote 8), we have found low rates of object-initial sentences overall (6% of 1510 sentences, with equal proportions in written and spoken samples), but extremely low rates with sentence-initial, genitive case-marked objects: only 4 sentences (0.3%) in our sample.

nouns may be interpreted as objects or possessors, and the likelihood of a reader or hearer considering each of these possible roles is influenced by factors such as animacy and number. Thus, when a case-marked noun is encountered without any linguistic or discourse context, the case-marker does not seem to provide specific information for mapping it directly to a grammatical function. The morphological case seems to function as a partial cue rather than a 'deterministic' marker.

The interpretations of grammatical roles triggered by partitive and genitive case point to information contributed by each case-marker: while partitive is strongly associated with argument functions (subject, object), genitive is used almost exclusively for non-argument functions, at least in the sentence-initial positions that we tested. Pinning down the more precise information these case-markers signal is left for further study.

In addition, we also find evidence which suggests that comprehenders prefer structurally simple structures over more complex ones. In earlier work, Bornkessel-Schlesewsky & Schlesewsky (2006) suggest that "[i]n the absence of explicit information to the contrary, the human language comprehension system assigns minimal structures." This kind of 'simplicity-based' processing system predicts that, whenever possible, participants should opt for intransitives over transitive clauses, since intransitives involve less structure. Indeed, we found that nominative nouns more often prompt intransitive sentences with SV order than transitive sentences with SVO order.

One important note is that the intransitive preference should not be attributed to participants being 'lazy' and writing short, simple continuations. On the contrary, our data contains many examples of long, complex continuations, as exemplified in (19)-(20).

(19) *Kingituse   leidmine   võttis   aega,   kuna   sünnipäevalapse*
     present.GEN.SG   finding   took   time   because forbirthdaychild.GEN.SG
     hobidest   oli   vähe   teada.
     hobby.PL.ELA   was   little   know
     'Finding a present took time, because little was known about the hobbies of the birthday child.'

(20) *Luuletuse   ettelugemine   läks   tal   paremini*
     poem.GEN.SG   reciting   went   3SG.ADE   better
     kui   selle   kirjutamine.
     than   this.GEN   writing
     'Reciting a poem went better for him/her than writing one.'

Moreover, effects of animacy on interpretations of case-marked nouns depend on the case-marker itself. On one hand, animacy seems to be a genuinely influential factor, involved in determining the functions of genitive nouns, independent of number. With nominative nouns, on the other hand, animacy has no effect. With partitive nouns, it is number rather than animacy which indicates the likelihood of certain grammatical roles over others. This raises the question of whether animacy should be treated as an independent factor in determining grammatical role assignment (see also Section 2.1; and Bornkessel-Schlesewsky & Schlesewsky 2009b). In our study, animacy affects the interpretation of partitive and genitive marked nouns, but not nominative. More importantly, animacy interacts with number in interpreting partitive nouns, but not genitive nouns. Thus, animacy interacts differently with different case-markers.

In sum, the results from our experiment suggest that the interpretation of nouns and their grammatical roles, even in the absence of any information from the verb or the syntactic structure of the sentence, is narrowed based on semantic cues – such as animacy and number – as well as morphological information, namely case-marking. Furthermore, our findings regarding transitivity provide evidence in favor of a cognitive preference for simple structures. Frequency is also shown to have an effect on grammatical role assignment. We are exploring this in more detail in further work.

References

Aissen, Judith. 2003. Differential object marking: iconicity vs. economy. *Natural Language and Linguistic Theory* 21. 435–483.

Altmann, Gerry T. & Kamide, Yuki. 1999. Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition* 73(3). 247–264.

Bamyaci, Elif, Häussler, Jana & Kabak, Baris. 2014. The interaction of animacy and number agreement: An experimental investigation. *Lingua* 148. 254–277.

Bornkessel, Ina & Schlesewsky, Matthias. 2006. The extended argument dependency model: A neurocognitive approach to sentence comprehension across languages. *Psychological review* 113(4). 787–821.

Bornkessel-Schlesewsky, Ina & Schlesewsky, Matthias 2009a. *Processing Syntax and Morphology: A Neurocognitive Perspective.* Oxford: Oxford University Press.

Bornkessel-Schlesewsky, Ina & Schlesewsky, Matthias. 2009b. The Role of Prominence Information in the Real-Time Comprehension of Transitive Constructions: A Cross-Linguistic Approach. *Language and Linguistics Compass* 3(1). 19–58.

Branigan, Holly P., Pickering, Martin J. & Tanaka, Mikihiro. 2008. Contributions of animacy to grammatical function assignment and word order during production. *Lingua* 118. 172–189. doi:10.1016/j.lingua.2007.02.003

Bresnan, Joan, Cueni, Anna, Nikitina, Tatiana & Baayen, Harald R. 2007. Predicting the dative alternation. *Cognitive Foundations of Interpretation.* (Royal Netherlands Academy of Arts and Sciences.) 69–94.

Bybee, Joan L. & Hopper, Paul J. (eds.). 2001. *Frequency and the emergence of linguistic structure.* Vol. 45. Amsterdam: John Benjamins Publishing.

Cann, Ronnie & Miljan, Merilin. 2012. Differential case-marking: Syntactic descriptions and pragmatic explanations. *The Linguistic Review* 29(4). 585–605.

Christianson, Kiel & Ferreira, Fernanda. 2005. Conceptual accessibility and sentence production in a free word order language (Odawa). *Cognition* 98(2). 105–135.

Comrie, Bernard. 1989. *Language Universals and Linguistic Typology*, 2nd edn. Oxford: Blackwell.

Corbett, Greville G. 2000. *Number.* Cambridge: Cambridge University Press.

Dahl, Östen. 2008. Animacy and egophoricity: Grammar, ontology and phylogeny. *Lingua* 118(2). 141–150.

Demiral, Sükrü B., Schlesewsky, Matthias & Bornkessel-Schlesewsky, Ina. 2008. On the universality of language comprehension strategies: Evidence from Turkish. *Cognition* 106. 484–500.

Erelt, Mati, Kasik, Reet, Metslang, Helle, Rajandi, Henno, Ross, Kristiina, Saari, Henn, Tael, Kaja & Vare, Silvi. 1993. *Eesti keele grammatika II* [The Grammar of the Estonian Language]. Tallinn: ETA Eesti Keele Instituut.

Erelt, Mati, Erelt, Tiiu & Ross, Kristiina. 2007. Eesti keele käsiraamat, 3rd edn. [The Handbook of the Estonian language]. Tallinn: Eesti Keele Sihtasutus.

Grewe, Tanja, Bornkessel, Ina, Zysset, Stefan, Wiese, Richard, von Cramon, D. Yves & Schlesewsky, Matthias. 2006. Linguistic prominence and Broca's area: The influence of animacy as a linearization principle. *NeuroImage* 32. 1395–1402.

Grewe, Tanja, Bornkessel-Schlesewsky, Ina, Zysset, Stefan, Wiese, Richard, von Cramon, D. Yves & Schlesewsky, Matthias. 2007. The role of the posterior superior temporal sulcus in the

processing of unmarked transitivity. *NeuroImage* 35. 343–352.

Gries, Stefan Thomas & Divjak, Dagmar (eds.). 2012. *Frequency effects in language learning and processing* (Vol. 244). Berlin: Walter de Gruyter.

Hale, John. 2001. The information conveyed by words in sentences. *Journal of Psycholinguistic Research* 32(2). 101-123.

de Hoop, Helen & de Swart, Peter (eds.) 2008. *Differential Subject Marking. Studies in Natural Language and Linguistic Theory.* Dordrecht: Springer.

Huumo, Tuomas & Lindström, Liina. 2014. Partitives across constructions: on the range of uses of the Finnish and Estonian "partitive subjects". In Luraghi, Silvia & Huumo, Tuomas (eds.). Partitive Cases and Related Categories. 153−176. Berlin: De Gruyter.

Kaiser, Elsi & Trueswell, John C. 2004. The role of discourse context in the processing of a flexible word order language. *Cognition 94*. 113–147.

Kaiser, Elsi, Miljan, Merilin & Vihman, Virve-Anneli (under review). Estonian speakers' interpretations of morphological case: Implications for Case/Agree. *Linguistische Arbeiten.* de Gruyter.

Kittilä, Seppo. 2005. Recipient-prominence vs. beneficiary-prominence. *Linguistic Typology* 9(2). 269–297.

Koptjevskaja-Tamm, Maria & Wälchli, Bernard. 2001. The circum-Baltic languages: An areal-typological approach. In Dahl, Östen & Koptjevskaja-Tamm, Maria (eds.). *The Circum-Baltic Languages: Typology and Contact.* Vol. 2, 615–750. Amsterdam: John Benjamins Publishing Co.

MacDonald, Maryellen C. 2013. How language production shapes language form and comprehension. *Frontiers in Psychology 4(226).* 1–16. doi: 10.3389/fpsyg.2013.00226

Lago, Sol, Shalom, Diego E., Sigman, Mariano, Lau, Ellen F. & Phillips, Colin. 2015. Agreement processes in Spanish comprehension. *Journal of Memory and Language* 82. 133–149.

Lamers, Monique J.A. 2005. The on-line resolution of subject-object ambiguities with and without case-marking in Dutch: Evidence from Event-related brain potentials. In Amberber, Mengistu & de Hoop, Helen (eds.). *Competition and Variation in Natural Languages: The Case for Case.* (Perspectives on Cognitive Science), 251–293. Oxford: Elsevier.

Lamers, Monique J.A. & de Hoop, Helen. 2004. Animacy information in human sentence processing: an incremental optimization of interpretation approach. In Christiansen, Henning, Skadhauge, Peter Rossen & Villadsen, Jorgen (eds.), *Constraint Solving and Language Processing* (Lecture Notes in Computer Science. Vol. 3438). Berlin, Heidelberg: Springer. 158–171.

Lamers, Monique J.A. & de Swart, Peter (eds.). 2012. *Case, Word Order, and Prominence: Interacting Cues in Language Production and Comprehension.* (Studies in Theoretical Psycholinguistics. Vol. 40) Dordrecht: Springer.

Levy, Roger. 2008. Expectation-based syntactic comprehension. *Cognition 106*(3). 1126–1177.

Lindström, Liina. 2004. Sõnajärg lause tuumargumentide eristajana eesti keeles. [Word order distinguishing between core arguments in Estonian]. In Lindström, Liina (ed.) *Lauseliikmeist eesti keeles* [Grammatical relations in Estonian]. (Tartu Ülikooli eesti keele õppetooli preprindid 1.), 40–49, Tartu: Tartu Ülikooli Kirjastus.

Mak, Willem M., Vonk, Wietske & Schriefers, Herbert. 2002. The influence of animacy on relative clause processing. *Journal of Memory and Language* 47. 50–68.

Mak, Willem M., Vonk, Wietske & Schriefers, Herbert. 2006. Animacy in processing relative clauses: The hikers that rocks crush. *Journal of Memory and Language*, 54, 466–490.

Metslang, Helena. 2013. *Grammatical relations in Estonian: subject, object, and beyond.* PhD thesis. University of Tartu.

Prat-Sala, Mercè & Branigan, Holly P. 2000. Discourse constraints on syntactic processing in language production: a cross-linguistic study in English and Spanish. *Journal of Memory and Language* 42. 168–182.

Rajandi, Henno & Metslang, Helle. 1979. *Määramata ja määratud objekt* [The Indefinite and definite object]. Tallinn: Valgus.

Roosmaa, Tiit, Koit, Mare, Muischnek, Kadri, Müürisep, Kaili, Puolakainen, Tiina & Uibo, Heli. 2003. Eesti keele arvutigrammatika: mis on tehtud ja kuidas edasi? *Keel ja Kirjandus* 3. 192–209.

Rosenbach, Anette. 2005. Animacy versus weight as determinants of grammatical variation in English. Language 81(3). 613–644.

Rosenbach, Anette. 2008. Animacy and grammatical variation – Findings from English genitive variation. *Lingua* 118. 151–171.

Schlesewsky, Matthias & Bornkessel, Ina. 2004. On incremental interpretation: degrees of meaning accessed during sentence comprehension. *Lingua* 114. 1213–1234.

Schlesewsky, Matthias & Bornkessel, Ina. 2006. Animacy in processing relative clauses: the hikers that rocks crush. *Journal of Memory and Language* 54. 466–490.

Snedeker, Jesse & Trueswell, John C. 2004. The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology*, 49(3). 238–299.

de Swart, Peter. 2005. Noun Phrase Resolution: The Correlation between Case and Ambiguity. In Amberber, Mengistu & de Hoop, Helen (eds.) *Competition and Variation in Natural Languages: The Case for Case.* (Perspectives on Cognitive Science), 205–222. Oxford: Elsevier.

Tael, Kaja. 1988. *Sõnajärjemallid eesti keeles (võrrelduna soome keelega)* [Word order patterns in Estonian (in comparison with Finnish)]. (Preprint, KKI-56.) Tallinn: Eesti Keele ja Kirjanduse Instituut.

Tamm, Anne. 2004. *Relations between Estonian verbs, aspect and case.* PhD thesis. Budapest: ELTE BTK.

Tanaka, Mikihiro, Branigan, Holly P. & Pickering, Martin J. 2005. The role of animacy in Japanese sentence production. Paper Presented at CUNY Conference, Tucson Arizona, USA.

Tanenhaus, Michael K., Spivey-Knowlton, Michael J., Eberhard, Kathleen M. & Sedivy, Julie C. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268. 1632–1634.

Traxler, Matthew J., Morris, Robin K. & Seely, Rachel E. 2002. Processing subject and object relative clauses: evidence from eye movements. *Journal of Memory and Language* 47. 69–90.

Traxler, Matthew J., Williams, Rihana S., Blozis, Shelley A. & Morris, Robin K. 2005. Working memory, animacy, and verb class in the processing of relative clauses. *Journal of Memory and Language* 53. 204–224.

Trueswell, John C., Tanenhaus, Michael K. & Kello, Christopher. 1993. Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory and Cognition* 19(3). 528–553.

Wu, Fuyun, Kaiser, Elsi & Andersen, Elaine. 2012. Animacy effects in Chinese relative clause processing. *Language and Cognitive Processes* 27(10). 1489–1524.

Merilin Miljan
Tallinn University, School of Humanities & University of Tartu, Institute of Estonian and General Linguistics
merilin.miljan@tlu.ee OR merilin.miljan@ut.ee

Elsi Kaiser
University of Southern California, Department of Linguistics
emkaiser@usc.edu

Virve-Anneli Vihman
University of Tartu, Institute of Estonian and General Linguistics
virve.vihman@ut.ee