Katalin Nagy C., Enikő Németh T. & Zsuzsanna Németh

Merging various data analysis techniques in pragmatics

Abstract

We aim to explore the merging of various data sources in two different disciplines of the study of language use, namely, synchronic and historical pragmatics. On the basis of two case studies taken from the actual research practice of these traditionally corpus-based fields, we demonstrate that corpus as a data source is of an integrated nature and research methodologies in pragmatic research have fuzzy boundaries. We attempt to show that even in a mainly corpus-based research practice various pieces of information gained from different sources are merged, consequently, corpus as a data source is not exclusive. We argue that it is complemented with the researchers' linguistic intuition, the results of earlier investigations, the theoretical framework we work in, inferences, etc. as further data sources.
Keywords: implicit arguments, context types in grammaticalisation, fuzzy boundaries of data sources, corpus data, linguistic intuition

1 Introduction

In the last decades metatheoretical studies have emphasised that the integration of data from various sources is useful and even necessary in linguistic research practice (cf. Lehmann 2004; Kepser & Reis 2005; Penke & Rosenbach 2004/2007; Kertész & Rákosi 2008, 2012, 2014). The different approaches to the study of language use assume that they collect and use data from well-distinguished sources compatible with their underlying theoretical considerations, such as linguistic intuition, written and spoken corpora, real and thought experiments, etc. However, the actual research practice often reveals

that there are fuzzy boundaries between these methods. In the present paper, we aim to examine this fuzziness in the research methodologies of two different disciplines of the study of language use, namely, synchronic and historical pragmatics. Although classic philosophical pragmatic studies dominantly applied data from intuition, introspection, and thought experiments, in the last few decades both synchronic and historical pragmatic research became corpus-based (corpus informed) or corpus-driven to a large extent in the sense that they mainly rely on the empirically observable occurrences of various linguistic and pragmatic units in spoken or written discourses.[1] However, data source in the corpus-based or corpus-driven research is more complex than it has been assumed previously.

To fulfil our aim, we will present two brief case studies in which methodological and metatheoretical considerations help to identify the different pieces of data merged during the research. Furthermore, we attempt to show that even in a mainly corpus-based research practice various pieces of information gained from different sources are merged, consequently, corpus as a data source is not exclusive. We argue that it is complemented with the researchers' linguistic intuition, the results of earlier investigation, the theoretical framework we work in, inferences, etc. as further data sources.

The organisation of the paper is as follows. After the introductory Section 1, in Section 2 we will briefly summarise the methodological and metatheoretical background underlying our hypothesis according to which it is not possible to draw clear boundaries between various data sources. In Sections 3 and 4, we will present two case studies to support our hypothesis. In Section 3, we will analyse implicit plural pronominal objects in Hungarian language use, and then, in Section 4, we will investigate the role of context types in semantic change of grammaticalisation. After presenting these case studies, in Section 5, we will summarise our results and conclusions.

---

[1]   Although in this paper we relied on data from various corpora, we did not aim to conduct a frequency-based statistical analysis. Consequently, our investigations were rather corpus-based than corpus-driven. For more details on the difference between corpus-based (corpus-informed) and corpus-driven investigation, see Tognini-Bonelli (2001).

2 Methodological and metatheoretical approaches to data and data sources in the study of language use

2.1 Traditional concepts of data and data sources

Jucker (2009: 1615-1619) discusses the "armchair", "field" and "laboratory" approaches as three different ways of doing pragmatic research.[2] The three types of data which these approaches provide are usually regarded as the three basic data types in the research of language use. Let us consider these three types of data in detail.

The first type of data includes data gained with the "armchair" method. Their source can be the researcher's intuition and introspection, but opinions and assessments from other speakers of a language or language variety can be elicited via the interview method (Jucker 2009: 1615).

The second type of data contains data gained with the "field method". According to Jucker (2009: 1615-1618), the field method is based on observing naturally occurring data, that is, utterances produced for communicative purposes outside the research project for which they are collected.[3]

And finally, the third type of data involves data gained with the "laboratory" method. The laboratory method (Jucker 2009: 1618-1619) includes different techniques to elicit certain utterances expected to be used in imaginary situations. This method enables researchers to control many different variables.

Although these types of data and data sources are widely accepted and applied in pragmatics, different approaches prefer one or another data source. At the same time, historical research obviously lacks some of these data types. Although researchers' intuition as a data source plays a role in historical research in order to judge the well-formedness of linguistic units of earlier language stages, researchers have to rely on what has been termed "substitute competence" (see Forgács 1993-1994), while native language speakers' linguistic intui-

---

[2]   These notions were introduced by Clark & Bangerter (2004: 25).

[3]   Jucker (2009: 1615-1618) enumerates several types of field method. Researchers can take notes of occurrences encountered in daily life ("notebook method"), collect occurrences from fictional material or from written documents ("philological method"), use transcriptions of actual conversations ("conversation analytical method"), and, finally, employ computerised search techniques on electronic corpora ("corpus method").

tions can be accessed only indirectly.[4] The philological and corpus methods can only be used when investigating language stages from which we have written documents. Other language stages can only be studied by relying on reconstructed data. Methods which require native language informants' contribution (i.e. the "notebook", the "conversation analytical", and the "laboratory" methods, see fn. 3) cannot be applied in historical linguistics. However, it is worth noting that data sources not available in the investigation of earlier language stages can still play a role in an indirect way, since argumentation in historical research can use information concerning the current language state.

Since both case studies presented here are corpus-based in the sense that they rely on the observable occurrences of various units in spoken or written discourses, some remarks on the notion of corpus data seem to be in order. Researchers in the wide domain of the study of language use usually share Lehmann's (2004: 201) opinion that data relying on corpora are the most reliable type of data. According to Lehmann, since corpus data are independent of the researcher, the research itself can be considered more objective. Francis (1992: 17) defines linguistic corpus as "a collection of texts assumed to be representative of a given language, dialect, or other subset of a language, to be used for linguistic analysis". The amount of texts contains a group of occurrences which represent the linguistic or pragmatic unit under study in the corpus. These occurrences are traditionally conceived of as data used in that particular research. However, if we use the term *data* in this sense, we cannot account for the practice that the researchers do not rely on the occurrences themselves, but the *statements* about them (cf. Kertész & Rákosi 2012: 170-171), that is, the relationship between occurrences and theoretical statements of a certain research is not explicit. In order to reveal this relationship, it is reasonable to apply a new concept of data.

---

[4]   For instance, through analysing grammars of the period at issue concerning normative language use, differences between original manuscripts and possible later duplicates or translations, remarks on the margins of historical documents, translations, etc.

2.2 *A new concept of data proposed by Kertész & Rákosi (2012)*

Kertész & Rákosi (2012) propose a new model of linguistic research, the central idea of which is that scientific theorising is a process of plausible argumentation based on only partially confirmed, i.e. plausible information and provides plausible results. Kertész & Rákosi (2012: 170-171) argue that the occurrences themselves cannot enter the argumentation process directly, only in form of statements. Kertész & Rákosi (2012: 169) define a datum as follows: "A datum is a statement with a positive plausibility value originating from a direct source".[5] According to this novel approach, data (plausible statements) combine information content and plausibility value, which is assigned to the information content on the basis of some direct source. These statements cannot be regarded as true or false, but as acceptable only to a certain extent (see Kertész & Rákosi 2012: 63). The plausibility of data is connected to the strength of acceptability of their source.

One of the advantages of Kertész & Rákosi's (2012) approach is that it can handle the traditional data types in a unified way. Data gained with the "armchair method", "field method" and "laboratory method" all can be regarded as plausible statements about occurrences originated from various sources.

Another advantage is that the model also enables a statement to be plausible on the basis of a source and implausible on the basis of another one at the same time (Kertész & Rákosi 2012: 169-184, 2014: 37-46),[6] therefore it is suitable to reveal inconsistencies emerging during the research.

Since both fields we work in are mainly corpus-based, we have to clarify the concept of corpus data we work with. We share Kertész & Rákosi's (2012: 173) view on corpus data, according to which corpus data are plausible statements about the presence of a linguistic unit in a corpus or some of its characteristics.

The question arises how data are used in the process of linguistic theorising and what the relationship is between data and hypotheses.

---

[5] According to Kertész & Rákosi's model, indirect sources are those from which one can get plausible statements by means of inferences, all other sources (e.g. intuition, corpus, and experiments) are considered direct.

[6] When a source supports a statement's negation, we can say that the statement is implausible on the basis of this source.

To answer this question, Kertész & Rákosi (2012: 178-184, 2014: 41-46) use the concept of evidence. Evidence is a relative notion in the model, and not an objective, given subcategory of data. Data function as evidence in the case when they "contribute to the judgement and comparison of the plausibility of rival hypotheses" (Kertész & Rákosi 2012: 178). If data from different sources function as evidence for a hypothesis, then higher plausibility value can be assigned to the hypothesis. On the basis of the above metatheoretical considerations, we can assume that the conscious integration of data from various data sources increases the plausibility and reduces the uncertainty resulted from relying on a single data source (Kertész & Rákosi 2012: 239).

In the next two sections, we will examine the merging of different data sources in synchronic and historical pragmatics. We will show that although these areas are mainly corpus-based, in the actual research practice they integrate data from different sources. However, the connections between these sources are not easy to reveal, since they can have fuzzy boundaries. Firstly, we present a case study in synchronic pragmatics.

3   Implicit plural pronominal objects in Hungarian language use

3.1   Aim

In this section we aim to examine a particular pragmatic phenomenon, namely, the occurrence of the implicit plural pronominal object arguments in Hungarian language use. To investigate this phenomenon and eliminate the inconsistency in the literature regarding their presence in Hungarian is only possible if we merge various data sources and assume fuzzy boundaries between them.

In the past two decades special attention has been devoted to the investigation of implicit verbal arguments in different languages in a complex approach which considers both grammatical and pragmatic (contextual) information interacting with each other (Cote 1996; García Velasco & Portero Muñoz 2002; Goldberg 2005; Németh T. 2010, 2017). Implicit arguments can be defined as arguments in lexical-semantic representations of verbs which are lexically unrealised, and whose implicit presence in utterances is attested by lexical-

semantic, grammatical, and/or pragmatic (contextual) evidence (Németh T. 2014, 2017).

In Hungarian language use the occurrence of verbs with implicit arguments can be licensed by lexical-semantic, grammatical, and pragmatic factors in the following three manners: (A) if some element of the lexical-semantic representation of a verb licenses the lexically unrealised occurrence of the argument, according to the principle of relevance (cf. *A mama főz [ételt]*[7] 'The grandmother is cooking [dish]'), (B) if the rest of the utterance, i.e. immediate context with its contextual factors including encyclopaedic pieces of information and grammatical requirements provides a relevant, typical interpretation (cf. *Rita adott a koldusnak [pénzt]* 'Rita gave [money] to the begger'), and, (C) if extending the immediate utterance context of the argument results in a relevant interpretation (cf. *Kihűltek a sütemények$_i$. Megettük pro$_{subj}$=[mi] pro$_{obj}$=[azokat$_i$].* 'The cookies got cool. We ate [them.ACC].').

Hungarian is a pro-drop language which licenses not only subject pro-drop but object pro-drop as well. Dropped objects are kinds of implicit direct object arguments which can be licensed and identified in the second (B) or the third (C) manner (Németh T. forthcoming).

Hungarian grammatical tradition as well as the current generative grammatical approaches agree that Hungarian transitive verbs can only be used with singular zero pronominal objects (Kugler 2000; É. Kiss 2002, 2012; see e.g. (2)). Their argumentation is mostly based on the sentence-level data coming from their own intuition. However, there are other Hungarian native speakers whose intuition does accept plural zero pronominal objects (see e.g. (1c)). Consequently, there is an inconsistency in the research regarding the acceptability of the plural zero pronominal objects. But if data from new data sources are taken into consideration and merged in the research, this inconsistency can be eliminated.

3.2 *Implicit pronominal objects in Hungarian language use*

The pronominal direct object can be left implicit in Hungarian language use, if it is expressed by the verbal inflection on the transitive verbs and/or can straightforwardly be identified in the context (see

---

[7] Implicit arguments are provided in square brackets.

e.g. (3)). In Hungarian language use both zero object anaphors and extralinguistically licensed object pro-drop can occur (see e.g. (7) and (12), respectively).

In Hungarian there are two types of conjugations, traditionally called indefinite and definite conjugations. Transitive verbs can be conjugated for both indefinite and definite forms, but intransitive verbs only for indefinite forms. If transitive verbs are used with an indefinite direct object, they need to be conjugated for the indefinite forms. If they are used with a definite direct object, they must be conjugated for the definite forms. The 1st and 2nd person dropped objects are indicated by the inflectional morphemes on the transitive verbs conjugated for the indefinite form; they do not occur with definite conjugations (É. Kiss 2012: 194). Let us consider (1a–c).[8]

(1)   a.   *Szeretsz/   szerettek   [engem/minket]?*
           love.INDEF.2SG   love.INDEF.2PL   me.ACC/us.ACC
           'Do you.SG.NOM/PL.NOM love [me/us]?'

       b.   *A   nagypapa   szeret   [engem/minket/*
           the   grandfather.NOM   loves.INDEF   me.ACC/us.ACC/
           téged/titeket].
           you.SING.ACC/you.PL.ACC
           'The grandfather loves [me/us/you.SG.ACC/you.PL.ACC].'

       c.   *Szeretlek   [téged/titeket].*
           love.1SG.2OBJ   you.SG.ACC/you.PL.ACC
           'I love [you.SG.ACC/you.PL.ACC].'

In (1a) the verb *szeret* 'love' is conjugated for the indefinite form, it agrees with the singular or plural 2nd person subject and the inflectional morpheme can refer to both the singular and plural 1st person dropped object indicating the actual speaker(s). In (1b) the verb *szeret* 'love' is also conjugated for the indefinite form but it has a singular or plural 3rd person subject and its inflectional morphemes can indicate both a singular and plural 1st or 2nd person dropped object, identifiable with the actual speaker(s) and hearer(s). If the subject is 1st person singular, the verbal suffix *-lak/-lek* can indicate both a singular

---

8   Abbreviations used in the glosses are provided at the end of the paper in the List of abbreviations.

and plural 2nd person dropped object which refer(s) to the actual hearer(s), cf. (1c).

3.3 *Hungarian grammatical tradition on implicit pronominal objects*

É. Kiss (2012: 193-195) assumes that Hungarian allows only singular null pronominal objects. If a transitive verb is conjugated for the indefinite form and occurs without an explicitly expressed 1st or 2nd person object, then a zero singular 1st or 2nd person object is supposed. So, according to É. Kiss's (2012: 193) claim, utterances in (1a−b) are only grammatical if they contain a singular null 1st or 2nd person object. However, in (1a) if the null 1st person pronominal object refers to the speaker and some other people, it definitely has a plural reading. Similarly, in (1b) if the null 1st and 2nd person objects refer to the speaker and some other people and to the hearer and some other people, respectively, they have a plural reading. The singular 1st and 2nd readings are only a default interpretation due to the lack of specific context. É. Kiss (2012: 195) especially highlights that the plural 2nd person object pronoun cannot be dropped with verbs conjugated for the *-lak/-lek* form. Therefore, plural reading in (1c) and in the second utterance in (2) are ungrammatical according to her intuition.

(2) *Ne   bújjatok   el!   Látlak   titeket/*
     not   hide.IMP.INDEF.2PL   PVB   see.1SG.2OBJ you.PL.ACC
     **pro.*
     [you.PL.ACC]
     'Don't hide! I can see you.PL'

Previous approaches to the objectless use of Hungarian transitive verbs share É. Kiss's (2012) opinion, i.e. they assume that plural 2nd person objects always have to be explicitly expressed with verbs conjugated for the *-lak/-lek* form (H. Molnár 1962: 157; Pete 1998: 140; Kugler 2000: 110). Thus, on the basis of their intuition as a reliable source, a statement that Hungarian transitive verbs cannot occur with zero plural pronominal objects seems to have a high plausibility value.

3.4   New data sources in the research on implicit plural pronominal objects

3.4.1   Data from intuition and thought experiments

In contrast to the Hungarian grammatical tradition, according to our intuition as well as that of other Hungarian native speakers (see Németh T. forthcoming), (1c) and the second utterance in (2) are also grammatical and acceptable when *szeretlek* 'I love you' and *látlak* 'I can see you' occur with an implicit plural 2nd person personal object pronoun [titeket 'you.PL.ACC']. However, it must be noted that there are Hungarian native speakers whose intuition considers plural reading in (1c) and in the second utterance in (2) questionable, i.e. these native speakers evaluate these occurrences as neither absolutely acceptable nor totally unacceptable. However, if we perform a thought experiment and situate these utterances in a particular context, the native speakers in question change their acceptability evaluations from questionable to acceptable, cf. (3) and (4), respectively.

(3)   (Children are playing hide-and-seek. The children whose turn it is to hide are hesitating where to hide, and so they are running out of time. The boy who finishes the counting says:)
Ne   bújjatok   el!   Látlak   [titeket].
not   hide.IMP.INDEF.2PL   PVB   see.1SG.2OBJ you.PL.ACC
'Don't hide! I can see [you.PL.ACC].'

(4)   (Grandchildren are behaving terribly; they are shouting and quarrelling with each other. The grandfather punishes them. The grandchildren become frightened and grow sad. After a while the grandfather says:)
Gyertek   ide gyorsan! Szeretlek   ám
come.IMP.INDEF.2PL   here quickly   love.1SG.2OBJ   really
[titeket],   nincs   semmi   probléma.
you.PL.ACC   not.is.INDEF nothing problem.NOM
'Come here quickly. I do love [you.PL.ACC], there is no problem.'

In (3) and (4) the contexts support the interpretation with zero plural 2nd person pronouns [titeket 'you.PL.ACC']. The implicit plural 2nd person subject indicated by the verbal inflection can serve as an antecedent for the zero objects of *látlak* 'I can see you' and *szeretlek* 'I love you' in the discourse context. However, it is worth emphasising that if we imagine the contexts described above without the first utterances,

Látlak [titeket] 'I can see you.PL.ACC' and *Szeretlek ám* [titeket] 'I really love you PL.ACC' still remain grammatical and acceptable with a zero plural 2nd object reading. Thus, the physical contextual information is strong enough to license the occurrence of the zero plural 2nd person object with the *-lak/-lek* inflection.

Let us modify the situation and utterances with 3rd person subjects. Cf. (5) and (6).

(5)   (Children are playing hide-and-seek. The children whose turn it is to hide are hesitating where to hide, and so they are running out of time. The boy who finishes the counting begins looking and can see other children not hiding. A girl realises that the boy who has begun looking can see them:)
Ne bújjatok el! Lát [titeket].
not hide.IMP.INDEF.2PL PVB sees.INDEF you.PL.ACC
'Don't hide! He can see [you.PL.ACC].'

(6)   (Children are behaving terribly; they are shouting at their mother. The grandfather punishes them. The children become frightened and grow sad. After a while the mother says:)
Szaladjatok oda gyorsan nagypapához!
run.IMP.INDEF.2PL there quickly grandfather.ALL
Szeret ám [titeket], nincs semmi
loves.INDEF really [you.PL.ACC] not.is.INDEF nothing
probléma.
problem.NOM
'Run there to the grandfather quickly. He does love [you.PL.ACC], there is no problem.'

The occurrence of implicit plural 2nd person object pronouns in (5) and (6) also attest that in Hungarian language use it is not only the singular 2nd person object pronoun which can be dropped. Furthermore, similar analysis can be provided in the other forms of indefinite conjugation (cf. Németh T. forthcoming).

The analyses of the utterances in (1)−(6), based on our own and other native speakers' intuition, as well as thought experiments, suggest with a high plausibility value that plural 2nd person pronominal objects can also be dropped in Hungarian language use if the particular contextual factors and/or anaphoric relations in subsequent utterances license it.

3.4.2 Corpus data

In order to eliminate the inconsistency concerning the occurrence of zero plural pronominal objects, let us introduce a new data source into the research, namely, the Hungarian National Corpus (corpus.nytud.hu/mnsz) and check whether there are occurrences of verbs conjugated for the *-lak/-lek* form with zero plural 2nd person objects. Cf. (7)–(8).

(7) *Eltévedt   madaraim$_i$,   látlak   [titeket$_i$],*
     PVB.lost   my birds.NOM   see.1SG.2OBJ you.PL.ACC
     látlak   [titeket$_i$].
     see.1SG.2OBJ   you.PL.ACC
     'My lost birds, I can see [you.PL.ACC], I can see [you.PL.ACC].'

(8) *Ági,   Margit,   Kati,   hol   vagytok [Ø$_{subj\ i}$]?*
     Ági.NOM   Margit.NOM   Kati.NOM   where   are.INDEF.2SG
     Szeretlek   [titeket$_i$]!
     love.1SG.2OBJ   you.PL.ACC
     'Ági, Margit, Kati, where are [you.PL.NOM]? I love [you.PL.ACC].'

The occurrence of *látlak* 'I can see you' in (7) and *szeretlek* 'I love you' in (8) with zero plural 2nd person objects can also be analysed as anaphoric null plural 2nd person pronominal objects which are coreferential with their coindexed antecedents in the discourse context. However, the zero plural 2nd person objects in (7)−(8) can also be analysed as zero exophoric objects since they also refer extralinguistically to the partners of communicators.

   Finally, let us examine the use and interpretation of the verbs with implicit 3rd person pronominal objects in Hungarian language use. É. Kiss (2012: 193) claims that if in a Hungarian sentence there is no overt object and the verb is conjugated for the definite form, a singular 3rd person pronominal object is assumed. Cf. É. Kiss's (2012: 194) example in (9).

(9) *Ismerem   pro$_{subj}$ =[én]   pro$_{obj}$ =[őt/azt].*
     know.DEF.1SG   I   her.ACC/him.ACC/it.ACC
     'I know [her.ACC/him.ACC/it.ACC]'.

According to É. Kiss (2012: 194), the plural pronominal object cannot be dropped, since the plural feature cannot be reconstructed from the

verbal suffix. However, she admits that there are plural zero 3rd person pronominal objects in Hungarian in the second conjuncts of coordinated sentences and in responses to *yes–no* questions. They are licensed when their antecedent is an object in the previous clause, but they are not anaphoric object *pro*; instead, their use involves VP-deletion. Cf. É. Kiss's (2012: 194) example in (10).

(10) *Az   ismerőseimet   keresem,   de   nem*
     the   my.acquaintances.ACC   seek.DEF.1SG   but   not
     találom   [VP 0]
     find.DEF.1SG
     'I am looking for my acquaintances but I cannot find [them.ACC].'

However, we can find occurrences of zero plural 3rd person pronominal objects when their antecedent in the context is not an object, and so VP-deletion cannot take place. Cf. (11).[9]

(11) *Kihűltek   a   süteményeki.   Megettük*
     PVB.got.cool.INDEF.3SG   the   cookies.NOM   PVB.ate.DEF.1PL
     prosubj=[mi]   proobj=[azokati].
     we   them.ACC
     'The cookies got cool. We ate [them.ACC].'

Although the verbal suffixes in (11) do not mark the plural feature either, as É. Kiss (2012: 194) claims, the verb *megettük* 'we PVB.ate', conjugated for the definite form occur with plural zero 3rd person pronominal objects. The particular context overrides the evaluation predicted by grammar, i.e. this anaphoric occurrence is acceptable, since it is licensed by particular contextual information. It is worth mentioning that exophoric plural zero 3rd person pronominal objects can also be licensed by contextual factors, as in (12).

(12) (The teacher writes a lot of four-digit numbers on the blackboard. Then she says:)
     Adjátok   prosubj=[ti]   össze   proobj=[azokat]!
     add.IMP.DEF.2PL   you   PVB   them.ACC
     'Add [them.ACC] up.'

---

9   Occurrences in (11) and (12) below have been collected via observation.

In (12) the verbal form *adjátok össze* 'add.PVB' conjugated for the definite form occurs with a plural zero 3rd person pronominal object which does not have any object antecedent in the previous utterances. In other words, it cannot be analysed as a case of VP-deletion either; instead, it has exophoric identification in the physical context. The deictic interpretation of the plural zero 3rd person pronominal object in (12) occurs similarly to the interpretation of the zero 1st and 2nd person pronominal objects discussed earlier, although the verb it occurs with is conjugated for the definite form, while the verbs with which zero 1st and 2nd person pronominal objects occur are conjugated for the indefinite forms.

Summarising the results, it can be concluded that in Hungarian language use transitive verbs can occur with zero plural pronominal objects both anaphorically and exophorically in all persons. The plausibility value of this concluding statement is much higher than the plausibility values of the inconsistent initial statements about the nonacceptability/acceptability of the occurrences of zero plural pronominal objects (see Sections 3.3 and 3.4.1). Data originated in the new data sources such as other native speakers' intuition, thought experiments and the Hungarian National Corpus have strengthened the plausibility value of the second initial statement according to which zero plural pronominal objects can occur in Hungarian language use. The case study has shown that the inconsistency in the literature regarding this problem can be eliminated only by introducing new data sources into the research and by merging various data types. The study has also highlighted how contextual factors in the thought experiments and corpus can influence the native speakers' intuition concerning the acceptability of implicit plural pronominal objects in Hungarian language use,[10] i.e. the merging of various data sources reveals fuzzy boundaries of data sources in the research practice.

In the next section, let us turn to our case study in historical pragmatics.

---

[10] This finding shows that these differences in the evaluation are not originated from dialectal diversity.

4 Context types in semantic change in grammaticalisation

4.1 Aim

In this Section, our aim is to show how corpus data are used in a particular area of historical pragmatics, namely, research on semantic change in grammaticalisation. As mentioned above in Section 2.1, in historical research we work with a restricted set of data sources, and the central role of corpus data is unquestionable. Fischer (2004: 730) holds that historical documents are the only firm source of knowledge for historical linguists. However, even in such a necessarily corpus-based research as historical research usually is, the source of data is more complex than it seems to be at first sight. In the actual research practice, corpus as data source is complemented with further sources and depending on the particular data sources we rely on, the analysis of corpus data can lead to different conclusions.

4.2 Context(s) in semantic change and contextual analysis in historical pragmatics

It is a widely accepted hypothesis in historical pragmatics that the meaning of linguistic constructions changes in – and as a result of – ordinary language use in context. Semantic change in grammaticalisation occurs in specific contexts and can be conceived of as a change of contexts in which a certain linguistic construction can be used. As a consequence, the description of changes in the semantic structure is closely linked to the description of contexts in which the grammaticalising linguistic item is used. To reconstruct changes in semantic structure and the spread of the innovation, we should examine the distribution of different context types among successive language stages: the occurrence or non-occurrence of a certain linguistic construction in a certain context type in historical documents can inform us about the process of semantic change. The nature of context types in which particular stages of semantic change processes can be detected and the importance of contextual analysis have been the topic of various papers on grammaticalisation (Diewald 2002; Heine 2002; Nagy C. 2014, forthcoming).

In order to investigate semantic change, we should be able to assign meanings to historical occurrences and to determine what kind of context we have in each case. That is to say, in each case the analysis

of semantic change presupposes the researcher's knowledge of the meaning of the particular linguistic unit in the historical context. In order to reveal what context type each and every occurrence belongs to, we need to check (i) whether the original, as well as the new interpretation is possible or not, (ii) whether it is the only possible reading, or only one of the accessible interpretations, and (iii) whether the new meaning is context dependent (i.e. emerges as an inference), or already independent of the context. We should minimise the subjectivity in identifying meanings and base our analysis on linguistic clues whenever possible. We need to bear in mind that ambiguity is always a necessary stage of a change in meaning. At the same time, we can base our analysis on contexts which favour or disallow one meaning as opposed to another, because historical occurrences in a type of context that excludes the original meaning can provide evidence for a previous semantic change.

4.3   Semantic change in the grammaticalisation of the Catalan "anar *'go' + infinitive": Two rival hypotheses*

The brief and simplified case study presented in this section concerns semantic change in the grammaticalisation of the Catalan "*anar* 'go' + infinitive[11]" construction. The combination of the verb meaning 'go' with an infinitive in its origin was a purposive construction, which referred to a motion with the aim of carrying out a certain act: 'go in order to do sg'. During the grammaticalisation of this periphrasis, the verb meaning 'go' lost its full lexical meaning of motion and became an auxiliary: the Catalan GO-construction, "*anar* 'go' + Inf", evolved into a preterit tense[12], and it is used to express a perfective past meaning in modern Catalan (cf. (13))[13].

(13)   *Pere II* **va**   **morir** *l'any 1285.*
                 AUX.PRS.3SG   die.INF
        'Pere II **died** in 1285.'

---

[11]   Infinitive: hereafter abbreviated as Inf.

[12]   The term *preterit tense* is used to express the combination of the perfective aspect and the past tense. In Catalan it is a synonym for perfective past.

[13]   In Section 4 we gloss only the relevant parts of the texts.

In the following, we present two main hypotheses on the formation of the Catalan GO-preterit and show how different the analysis of the same corpus data can be depending on which hypothesis we accept. The auxiliary forms of the perfective past "*anar* + Inf" in modern Catalan formally are in the present tense. Therefore, it is not surprising that some authors base their analyses on medieval present tense occurrences and derive the current Catalan "*anar* + Inf" construction from them. These authors regard relevant present tense occurrences of the verb *anar* as instances of a historical present usage, so I will refer to their approach as the *historical present hypothesis* (cf. Colon 1959/1978, 1976/1978; Pérez Saldanya 1996; Pérez Saldanya & Hualde 2003).

Another group of authors rejects the supposition that Catalan narrative texts containing early occurrences of the periphrasis can be characterised by the use of historical present. The most prominent representative of this approach is Juge (2002, 2006, 2008), who claims that forms in the current paradigm of the Catalan perfective past "*anar* + Inf" in the present tense resulted from analogy with other constructions which have the auxiliary in the present tense. He assumes that this process only took place after the past meaning of the whole construction had already consolidated. Consequently, the morphologically present forms are not examples of the historical present, but fully grammaticalised forms. Juge bases his hypothesis on a thorough morphological analysis of occurrences and on the observation that semantic change has already occurred in some examples with the verb *anar* in the preterit. I will therefore use the term *preterit hypothesis* to refer to this account, opposing it to the historical present hypothesis.

4.4 Context types in the grammaticalisation of the Catalan GO-construction

In a previous study, Nagy C. (2015) argued in favour of the preterit hypothesis through providing a pragmatic analysis of the process, relying on a detailed contextual analysis of occurrences found in a historical corpus of Catalan texts from the 13th to the 16th centuries. Various types of context have been differentiated, which, as they can be linked to successive stages of grammaticalisation, have helped to reveal the process of semantic change. When interpreting occurrences, Nagy C. followed the methodological principle according to which, if nothing

in the context contradicts it, we should assume the original lexical meaning of *anar* 'go'. Occurrences interpretable literally can also be classified into different subtypes. Historical data have suggested an ongoing increase of occurrences over time where the lexical meaning can be excluded. Let us see all these relevant context types.

(i) In some contexts the motion covers long distances and leads to a different scene. The continuation of the story facilitates an inference to the effect that the event indicated as the aim of the motion was actually performed. Although this piece of information is not explicitly expressed, it had to be inferred in order to have a coherent discourse. Observe the following example.

(14) *E no bastà açò, que enans* **anaren**   **combatre** *Nicòtena,*
                                go.PST.3PL   combat.INF

      e preseren-la;
      'And that was not enough; rather they **went to attack** Nicótena and took it;' (Munt I 102,34-35)[14]

In (14) the motion meaning is obvious, but the actual performance of the act described as the aim of this motion is not explicitly expressed. However, in order to have a coherent story, we assume the following events when interpreting the discourse in (14): they went to Nicótena with the aim of attacking it, they attacked it and finally, they took the city. The frequent use of the construction in this type of context might have established a strong associative link between the use of the periphrasis and the inferential content of perfectivity.

(ii) In another group of occurrences allowing for the motion reading, the movement is minimal and presumably performed within a very limited space, and sometimes the question arises as to whether it is

---

[14] The sources of the examples are abbreviated at the end of each example, with page and/or line numbers (full details of sources can be found at the end of the paper in Historical sources). The English version of Muntaner's Chronicle by Goodenough (2000) has been used to translate some of the historical text fragments. However, given that it is a literary translation, sometimes it was necessary to modify it for the linguistic purposes of the present discussion. For example, Goodenough (2000) translates the occurrence *anaren combatre* in (14) by the paratactic sequence "they went and attacked". This translation has not been adopted because it does not reflect the idea that we are dealing with an infinitival construction. As for city names, the spelling of the original text is used in the translations. Bibliographic data for the translation are given in Historical sources.

reasonable to assume any movement at all. In these contexts the aim of the narrator could be to achieve dynamism by the use of the motion verb and not to describe spatial movement at all. This interpretation appears in previous literature too, not only for Catalan, but for French as well, where the periphrasis is characterised as a certain kind of stylistic tool (cf. Colon 1976/1978: 144). The frequent use of "*anar* + Inf" in this type of context can also promote the semantic change of the construction. Consider the example in (15).

(15)   *E lo senyor rei saltà dins, avant, qui era jove e trempat,*
     *e **va-li**   tal **donar**   per mig del cap, de l'espaa, que*
       go.PRS.3SG   give.INF
     el capmall que portava no li valc un diner, que entrò
     *en les dents lo fenè; e   puis   **va**-li   **trer***
                           go.PRS.3SG   pull-out.INF
     *l'espaa del cap, e **va**   'n   **ferir**   altre, que*
                  go.PRS.3SG   attack.INF
     el braç ab tot lo muscle n'avallà en terra.
     'And the Lord King, who was young and spirited, advanced and **gave (=[15] goes to give) him such a blow** with his sword on the middle of his head, that the cap of mail he was wearing was of no use to him, for he was split open to the teeth. Then the King **pulled (= ?goes to pull)** the sword **out** of this man's head, and **attacked (= goes to attack)** another, whose arm, with the whole shoulder, fell to the ground.' (Munt II 50,3-8)

The fragment in (15) illustrates the typical early context of use of the periphrasis: a battle scene with fast successive movements within a relatively limited space. The first clause of the quotation describes a motion event that leads to the location of the battle: the king 'advanced' (*saltà dins*), 'forward' (*avant*), to the battleground, and after that we may assume short-range dynamic movements within a limited area of the battleground. The first occurrence of "*anar* + Inf" *va-li tal donar* '(he) goes to give him such a blow' allows the motion-then-action reading, but only if we interpret this verbal form as historical present. In contrast, this reading must be excluded in the case of the following occurrence of "*anar* + Inf" (*va-li trer* 'pulled out'), or at least it would be strange: after giving a big blow with his sword to a man,

---

[15] The symbol = is used to indicate literal meaning, while the symbol *?* indicates strange/nonsensical readings.

the king pulled the sword out of his head. These two movements must have taken place one after the other in quick succession, which does not imply any movement in space. Then the king turns to attack another man: in this case it is possible that he took some steps in the battlefield (if we interpret the verbal form as a historical present), although it was not absolutely necessary, because participants in a battle scene are supposed to be close to each other. This scene, if imagined as a dynamic battle scene, does not suppose long movements, although small movements cannot be excluded in two of the three occurrences. However, even in these cases the motion reading is only possible if we accept the historical present hypothesis. If we accept the preterit hypothesis, we should interpret these occurrences as fully grammaticalised forms.

(iii) Finally, at the conventionalisation stage, the new meaning no longer needs to be supported by the context. Once the target meaning has been conventionalised, the construction can appear in new contexts, as in (16).

(16) *Mas con foren prop d'Agda, noves los* **van**   **venir**
                                               AUX.PRS.3PL come.INF
      con havia pres, lo dia passat, a aquells de Besers
      'But when they came near Agda, news **arrived** (= **?go to come**) that they of Besers had been taken on the previous day.' (Munt II 10,36-37)

An advanced stage of grammaticalisation can be identified in (16), because the originally purposive construction "'go' + Inf" co-occurs with a subject that is incompatible with the source meaning of 'go'. The motion-with-intention meaning excludes inanimate subjects, thus occurrences as the one in (16) suggest that semantic change has already taken place. The combination of *anar* with the infinitive *venir* 'come' as in (16) is itself a contradiction due to the deictic meanings of 'go' and 'come', since the literal interpretation would yield the strange and nonsensical reading ?'the news come/are coming to go'. We do not need the broader context anymore to find out that this is not the literal meaning, i.e. this example is representative of the conventionalisation stage. The combination *van venir* can be analysed as a fully grammaticalised form conveying a preterit meaning independently of whether we accept the historical present or preterit hypothesis.

4.5 Different analyses of corpus data

Analysis of the distribution of these context types over time has led to the finding that the number of contexts where the 'motion' meaning is possible decreases over the period examined, while the number of contexts where it can be excluded increases (cf. Nagy C. 2015). Tables 1 and 2 show the assumed context types according to the tense of the finite verb in four texts of the historical corpus investigated.[16]

	Jau	Desc	Munt	Per	**total**
motion is possible	15 (~78.94%)	60 (~86.95%)	62 (~87.32%)	0	137
short-range movement	0	4 (~5.79%)	4 (~5.63%)	0	8
motion is not possible	4 (~21.05%)	5 (~7.24%)	5 (~7.04%)	1	15
total	**19**	**69**	**71**	**1**	**160**

Table 1: Distribution of context types in preterit tense occurrences of "anar + Inf"[17]

	Jau	Desc	Munt	Per	**total**
motion is possible	2	15 (~55.55%)	65 (41.4%)	1 (~7.69%)	83
short-range movement	0	3 (~11.11%)	26 (16.56%)	0	29
motion is not possible	0	9 (~33.33%)	66 (42.03%)	12 (~92.3%)	87
total	**2**	**27**	**157**	**13**	**199**

Table 2: Distribution of context types in present tense occurrences of "anar + Inf"

---

[16] As mentioned before, the present investigation did not require a frequency-based statistical analysis, it was rather corpus-based than corpus-driven (cf. Tognini-Bonelli 2001).

[17] The writing of the first text (Jau) can be dated between 1229 and 1276. However, the surviving manuscripts can be dated to a later period, the oldest one being from 1343. The date of composition of Desclot's chronicle (Desc) is not certain, but we can suppose that it must have been written between 1283 and 1289. Muntaner's chronicle (Munt) was written from 1325 to 1332. The last text (Per) is datable to the second half of the 14th and the early 15th centuries. For references to the texts see Historical sources.

In Table 1 it is worth observing that also in the case of the preterit examples there are contexts where the motion reading is impossible, although we can observe a reduction in their proportion. The reason for this finding might be that present tense occurrences became able to convey such readings. Table 2 reflects that present tense occurrences gradually became associated with the context type that excludes the source meaning of *anar* to a greater extent.

Now, let us compare how these pieces of data can be analysed according to the two hypotheses presented above. Historical present hypothesis accounts for the semantic change taking into consideration only the present tense occurrences. If we accept it, we will have two important consequences. Firstly, each and every piece of data included in Table 1 has to be ignored, because they are considered to have no relevance from the point of view of semantic change. Secondly, we have to examine the context of each and every occurrence in present tense in order to check whether they can convey motion meaning or not, that is, which context types they belong to. The path of semantic change will be delineated by the interpretation and classification of occurrences in present tense (cf. Table 2).

Preterit hypothesis accounts for the semantic change taking into consideration also the past tense occurrences. If we opt for it, we have to consider the contexts of preterit occurrences and classify them into different types. In contrast, occurrences in present tense must be interpreted as fully grammaticalised forms, in other words, they all will be included in the 'motion is not possible' row of Table 1. The result is shown in Table 3.

	Jau	Desc	Munt	Per	**total**
motion is possible	15 (~71,4%)	60 (62,5%)	62 (~27,2%)	0	137
short-range movement	0	4 (~4,16%)	4 (~1,75%)	0	8
motion is not possible	6 (~28,6%)	32 (~33,33%)	162 (~71,05%)	14 (100%)	214
total	**21**	**96**	**228**	**14**	**359**

Table 3: Distribution of context types in occurrences of "anar + Inf" according to the preterit hypothesis

When interpreting the occurrences this way, the tendencies mentioned above are more impressive. Let us take Muntaner's *Chronicle* as an example. If we reject the historical present hypothesis, 162 of

the occurrences in this text will be interpreted as grammaticalised forms, and if we accept it, only 71 of them. Relevant occurrences found in other texts of the corpus also would be analysed differently depending on the two hypotheses. As a consequence, two different paths of semantic change can be outlined. According to the preterit hypothesis, occurrences in the earliest text examined already suggest that the semantic change has started. However, every occurrence in this text considered to be relevant according to the historical present hypothesis conveys the source meaning. As a consequence, these two different ways of interpreting the same occurrences affect the dating of semantic change as well.

In this Section, we have argued that an inconsistency can emerge even if relying on the same occurrences found in the same historical texts. In this case study we have shown that different interpretations can follow from different theoretical stances. However, the consideration of further factors would lead to an even more complex picture. Other relevant factors, among others, could be the following: (i) how sensitive the contextual analysis is, i.e. how many relevant context types are hypothesised, (ii) the researchers' intuition, i.e. how they interpret each and every occurrence and how they classify them into context types, and, (iii) what kind of texts they use when searching for relevant occurrences, etc. The analysis of historical pragmatic research practice in using and interpreting corpus data has revealed the integrated nature of historical data. In order to make historical occurrences usable in linguistic research, we have to add other data from further sources, which shows how fuzzy the boundaries of data sources are in the actual research practice.

5 Conclusion

In this paper, on the basis of two case studies taken from the actual research practice we have shown that in the synchronic and historical pragmatic research various pieces of information gained from different sources are merged.

In our first, synchronic pragmatics case study, we have attested that the inconsistency in the literature regarding the nonacceptability/acceptability of the occurrences of zero plural pronominal objects in Hungarian language use can be eliminated only by introducing new data sources into the research and merging various data types. Data originated in new data sources such as other native speakers'

intuition, thought experiments and the Hungarian National Corpus have strengthened the plausibility value of the statement that zero plural pronominal objects can occur in Hungarian language use. The study has also highlighted how contextual factors in the thought experiments and corpus can influence the native speakers' intuition concerning the acceptability of implicit plural pronominal objects in Hungarian language use.

In our second, historical pragmatics case study, we have presented how differently corpus occurrences are analysed according to two rival hypotheses on the semantic change in grammaticalisation of the Catalan "*anar* 'go' + Inf". We have argued that an inconsistency can emerge even if relying on the same occurrences found in the same historical texts. Since in the actual research practice corpus as data source has to be complemented with further sources, the analysis of corpus occurrences can lead to different conclusions depending on the particular data sources we add. These different ways of interpreting the same occurrences outline different paths of semantic change and also affect the dating of semantic change.

From the metatheoretical point of view, both case studies have led to the same conclusion that corpus as a data source is of an integrated nature and the various research methodologies have fuzzy boundaries. We have argued that in the mainly corpus-based synchronic and historical pragmatic research corpus as a data source is not exclusive, but it is complemented with the researchers' linguistic intuition, the results of earlier investigation, the theoretical framework we work in, inferences, etc. as further data sources. A similar metatheoretical analysis would also be possible and desirable on research based on other data collecting methods.

List of abbreviations

1 = first person
2 = second person
3 = third person
ACC = accusative
ALL = allative
AUX = auxiliary
DEF = definite conjugation
IMP = imperative
INDEF = indefinite conjugation

INF = infinitive
NOM = nominative
OBJ = object
PL = plural
PST = past
PRS = present
PVB = preverb
SG = singular
SUBJ = subject

Historical sources

Jau: *Llibre dels fets del rei en Jaume* ['The Book of Deeds of James I of Aragon']. Vols 1–2. Barcelona: Barcino. (1991).

Desc: B. Desclot. *Crònica* ['Chronicle']. Vols. 1–5. Barcelona: Edicions 62. (1990).

Munt: R. Muntaner. *Crònica* ['Chronicle']. Vols. 1–2. Barcelona: Edicions 62. (1979/1990-1991). English version: Ramon Muntaner, *Chronicle*. Translated by Lady Goodenough. Cambridge, Ontario: In Parentheses Publications. (2000). Available at: http://www.yorku.ca/inpar/muntaner_goodenough.pdf.

Per: R. de Perellós: Viatge del vescomte Ramon de Perellós i de Roda fet al Purgatori nomenat de Sant Patrici ['The Journey of Viscount Ramon de Perellós i de Roda to Saint Patrick's Purgatory']. In *Novel·les amoroses i morals*. Barcelona: Edicions 62. (1982).

References

Clark, H.H. & Bangerter, A. (2004): Changing ideas about reference. In: Noveck, I.A. & Sperber, D. (eds.): *Experimental Pragmatics*. Houndmills: Palgrave Macmillan, 25-49.

Colon, G. (1959/1978): El perfet perifràstic català "va + infinitiu". In: Colon, G. (ed.): *La llengua catalana en els seus textos, 2. vol.* Barcelona: Curial, 119-130.

Colon, G. (1976/1978): Sobre el perfet perifràstic 'vado + infinitiu', en català, en provençal i en francès. In: Colon, G. (ed.): *La llengua catalana en els seus textos, 2. vol.* Barcelona: Curial, 131-174.

Cote, Sh.A. (1996): *Grammatical and Discourse Properties of Null Arguments in English.* PhD dissertation. Philadelphia: University of Pennsylvania.

Diewald, G. (2002): A model for relevant types of contexts in grammaticalization. In: Wischer, I. & Diewald, G. (eds.): *New reflections on grammaticalization.* Amsterdam: John Benjamins, 103-120.

É. Kiss, K. (2002): *The Syntax of Hungarian.* Cambridge: Cambridge University Press.

É. Kiss, K. (2012): Null pronominal objects in Hungarian: a case of exaptation. *Acta Linguistica Hafniensia* 44, 192-206.

Fischer, O. (2004): What counts as evidence in historical linguistics? *Studies in Language* 28, 710-740.

Forgács, T. (1993-1994): Zárt korpuszok és pótkompetencia [Closed corpora and substitute competence]. *Néprajz és Nyelvtudomány* 35, 17-23.

Francis, W.N. (1992): Language corpora B. C. In: Svartvik, J. (ed.): *Directions in Corpus Linguistics*. Berlin: Mouton de Gruyter, 17-32.

García Velasco, D. & Portero Muñoz, C. (2002): *Understood Objects in Functional Grammar* (Working papers in functional grammar 76). Amsterdam: University of Amsterdam.

Goldberg, A. (2005): Argument realization. The role of constructions, lexical semantics and discourse factors. In: Östman, J.-O. & Fried, M. (eds.): *Construction Grammars: Cognitive Grounding and Theoretical Extensions*. Amsterdam: John Benjamins, 17-43.

H. Molnár, I. (1962): A tárgy [Direct object]. In: Tompa, J. (szerk.): *A mai magyar nyelv rendszere. Leíró nyelvtan 2. Mondattan* [The Structure of Modern Hungarian Language. A Descriptive Grammar. Vol 2. Syntax]. Budapest: Akadémiai Kiadó, 147-161.

Heine, B. (2002): On the role of context in grammaticalization. In: Wischer, I. & Diewald, G. (eds.): *New reflections on grammaticalization.* Amsterdam: John Benjamins, 83-102.

Jucker, A.H. (2009): Speech act research between armchair, field and laboratory. The case of compliments. *Journal of Pragmatics* 41, 1611-1635.

Juge, M.L. (2002): *Tense and aspect in periphrastic pasts: evidence from Iberian Romance.* PhD Thesis. University of California-Berkeley.

Juge, M.L. (2006): Morphological factors in the grammaticalisation of the Catalan "go" past. *Diachronica* 23, 313-339.

Juge, M.L. (2008): Narrative and the Catalan GO-past. *Folia Linguistica historica* 29, 27-56.

Kepser, S. & Reis, M. (eds.) (2005): *Linguistic evidence. Empirical, theoretical and computational perspectives* (Studies in Generative Grammar 85). Berlin/New York: Mouton de Gruyter.

Kertész, A. & Rákosi, Cs. (2008): Introduction: The problem of data and evidence in theoretical linguistics. In: Kertész, A. & Rákosi, Cs. (eds.): *New Approaches to Linguistic Evidence. Pilot Studies / Neue Ansätze zu Linguistischer Evidenz. Pilotstudien.* Frankfurt am Main: Lang, 9-19.

Kertész, A. & Rákosi, Cs. (2012): *Data and Evidence in Linguistics: A Plausible Argumentation Model.* Cambridge: Cambridge University Press.

Kertész, A. & Rákosi, Cs. (2014): *The Evidential Basis of Linguistic Argumentation.* Amsterdam, Philadelphia: John Benjamins.

Kugler, N. (2000): Az igeragozás [The conjugation]. In: Keszler, B. (szerk.): *Magyar grammatika* [Hungarian Grammar]. Budapest: Nemzeti Tankönyvkiadó, 104-126.

Lehmann, Ch. (2004): Data in linguistics. *The Linguistic Review* 21, 175-210.

Nagy C., K. (2014): Kontextus(ok) a grammatikalizációs jelentésváltozás irodalmában [Context(s) in the literature on semantic change in grammaticalisation]. *Jelentés és Nyelvhasználat* 1, 11-27.

Nagy C., K. (2015): *Adatok és érvelés a történeti pragmatikában – A katalán* anar + *főnévi igenév szerkezet grammatikalizációja* (Pragmatika 2) [Data and argumentation in historical pragmatics. Grammaticalisation of the Catalan *anar* + infinitive construction]. Budapest: Loisir.

Nagy C., K. (forthcoming): *Data and argumentation in historical pragmatics. Grammaticalisation of a Catalan motion verb construction*. Sheffield: Equinox.

Németh T., E. (2010): How lexical-semantic factors influence the verbs' occurrence with implicit direct object arguments in Hungarian. In: Németh T., E. & Bibok, K. (eds.): *The Role of Data at the Semantics–Pragmatics Interface*. Berlin: Mouton de Gruyter, 305-348.

Németh T., E. (2014): Hungarian verbs of natural phenomena with explicit and implicit subject arguments: their use and occurrence in the light of data. In: Kertész, A. & Rákosi, Cs. (eds.): *The Evidential Basis of Linguistic Argumentation*. Amsterdam: John Benjamins, 103-132.

Németh T., E. (2017): Theoretical and methodological issues in the research into implicit arguments in Hungarian. In: Assimakopoulos, S. (ed.): *Pragmatics at Its Interfaces*. Berlin: Mouton de Gruyter, 149-174.

Németh T., E. (forthcoming): *Implicit Subject and Direct Object Arguments in Hungarian Language Use: Grammar and Pragmatics Interacting*. Sheffield: Equinox.

Penke, M. & Rosenbach, A. (eds.) (2004/2007): *What Counts as Evidence in Linguistics. The Case of Innateness* (Benjamins Current Topics 7). Amsterdam, Philadelphia: Benjamins.

Pérez Saldanya, M. (1996): Gramaticalització i reanàlisi: el cas del perfet perifràstic en català [Grammaticalisation and reanalysis: the case of the Catalan periphrastic past tense]. In: Schönberger, A. & Stegmann, T. D. (eds.): *Actes del desè col·loqui internacional de llengua i literatura catalanes, 3. vol.* Barcelona: Publicacions de l'Abadia de Montserrat, 71-107.

Pérez Saldanya, M. & Hualde, J.I. (2003): On the origin and evolution of the Catalan periphrastic preterit. In: Pusch, C.D. & Wesch, A. (eds.): *Verbalperiphrasen in den (ibero-)romanischen Sprachen*. Hamburg: Helmut Buske, 47-60.

Pete, I. (1998): A magyar igeragozás típusai [Types of Hungarian conjugation]. In: Büky, L. & Maleczki, M. (eds.): *A mai magyar nyelv leírásának újabb módszerei 3.* [Current Methods in the Description of the Present-Day Hungarian 3]. Szeged: JATE, 133-148.

Tognini-Bonelli, E. (2001): *Corpus Linguistics at Work* (Studies in Corpus Linguistics 6). Amsterdam, Philadelphia: John Benjamins.

Dr. Katalin Nagy C.
MTA-DE-SZTE Research Group for Theoretical Linguistics
Egyetem u. 2.
H-6722 Szeged
nagykati@hist.u-szeged.hu

Prof. Dr. Enikő Németh T.
University of Szeged, Department of General Linguistics
MTA-DE-SZTE Research Group for Theoretical Linguistics
Egyetem u. 2.
H-6722 Szeged
nemethen@hung.u-szeged.hu

Dr. Zsuzsanna Németh
MTA-DE-SZTE Research Group for Theoretical Linguistics
Egyetem u. 2.
H-6722 Szeged
nemethzs7@gmail.com