Theoretical Linguistics Programme, Budapest University (ELTE)

DOCSYMP:

GRADUATE STUDENTS' FIRST LINGUISTICS SYMPOSIUM

7. June 1996, Budapest

—Proceedings—

Research Institute for Linguistics, Hungarian Academy of Sciences

Working Papers in the Theory of Grammar, Vol.4, No.3
Received May 1997

Working Papers in the Theory of Grammar, Vol. 4, No. 3

DOCSYMP:

GRADUATE STUDENTS' FIRST LINGUISTICS SYMPOSIUM

7. June 1996, Budapest

—Proceedings—

Publication of this volume has been supported by *ELSNET*

Theoretical Linguistics Programme, Budapest University (ELTE)
Research Institute for Linguistics, Hungarian Academy of Sciences

Budapest I., P.O.B. 19, H-1250 Hungary
Telephone: (+36-1) 175 8285; Fax: (+36-1) 212 2050

Theoretical Linguistics Programme, Budapest University (ELTE)

DOCSYMP:

GRADUATE STUDENTS' FIRST LINGUISTICS SYMPOSIUM

7. June 1996, Budapest

—Proceedings—

Research Institute for Linguistics, Hungarian Academy of Sciences

Working Papers in the Theory of Grammar, Vol.4, No.3
Received May 1997

Foreword

This volume contains the Proceedings of the First Graduate Students' Linguistics Symposium, which took place on June the 7th, 1996 in Budapest. It was organized by the Theoretical Linguistics Programme at ELTE University (Budapest). As it was first in its kind some background information may be in order.

The conference was organized with the aim of providing a meeting-ground for graduate students and advanced undergraduates from newly established or reorganized Linguistics departments and doctoral programmes in Hungary. Accordingly, the papers in this volume reflect the quality of teaching and research at comparatively young institutions. The conference was open to *any* approach to language that involved some theoretically-minded investigation of data. In this light the papers in this volume, along with two presentations on psycholinguistics and neurolinguistics that appeared in other publications show remarkable consistency by their adherence to the generative paradigm in the broadest sense possible.

Papers in this volume follow the standard division of labour between phonology, syntax and semantics. Two presentations at the conference were on psycholinguistics; they are not included in this volume. One of them, by Katalin Kiss, is included in a special issue on psycholinguistics of *Acta Linguistica Hungarica*. The contribution of Péter Rebrus on Hungarian default vowels appeared as a joint paper with Krisztina Polgárdi in the book series *Approaches to Hungarian*. Huba Bartos' paper is a minimalist analysis of Hungarian (Definite) object agreement and the syntax-semantics interface. István Bujdosó defines Hungarian negation in Michael Bródy's Focus Phrase model and analyzes the interaction of negation and Focus. András Cser's contribution is a reconstruction of Latin syllable structure. Anikó Csirmaz analyzes Swahili possessive constructions and their agreement patterns in a minimalist framework. Kálmán Dudás examines impersonal constructions: his conclusions on Hungarian are backed with Latin and German data. Tibor Szécsényi offers an HPSG analysis of Hungarian infinitival constructions; he derives word order in these constructions from a merge of the SUBCAT list of the matrix verb and the infinitive. Péter Szigetvári's concern is the exact status of affricates; he backs his analysis with a wealth of Hungarian and Polish examples. Viktor Trón uses dynamic Kripke frames to handle English and Hungarian temporal dependencies that have so far been analysed with syntactic tools. Károly Varasdi explores the algebraic properties of connectives and negation in Dynamic Logic. Papers in this volume appear as the authors prepared them. Copyright resides with the individual authors.

The conference (nicknamed DOCSYMP) was made possible thanks to the generous help of several personalities and organisations. Funds from *ELSNET* were sufficient to cover the costs of the conference and the publication of this volume; the generosity of *ELSNET* is hereby gratefully acknowledged. From the Theoretical Linguistics Programme and the Research Institute for Linguistics of the Hungarian Academy of Sciences we owe thanks to professors Ferenc Kiefer, Zoltán Bánréti and László Kálmán, and to Mária Kováts, Huba Bartos, Gréte Dalmi, Viktor Trón, Károly Varasdi and Kinga Gárdai for their patience and assistance. We also wish to thank Mária Hanzséros, Ádám Nádasdy, Mark Newson, Péter Szigetvári and László Varga from the English Department of ELTE.

Budapest, May 1997

Ágnes Bende-Farkas

TABLE OF CONTENTS

Object Agreement Licensing in Hungarian

Huba Bartos

0. Introduction

Hungarian displays two verbal agreement paradigms, traditionally referred to as 'subjective' and 'objective' inflection ('alanyi ragozás' and 'tárgyas ragozás', respectively, in Hungarian). In very general terms, intransitive verbs are invariably affixed with the subjective endings, while in the case of transitive verbs, the choice depends on some property of the object. The fundamental nature of this decisive factor is the main topic of the present paper. I will argue that all the previous accounts of the nature of this 'object agreement' are unable to cover all the cases involved, because they all fail to recognize the precise properties that condition the choice between the paradigms. I will therefore propose a new criterion for the distinction between nominal phrases that trigger objective agreement, and those that do not. At the heart of my suggestion lies the assumption that nominal phrases are not uniform categorially: some project a DP-layer, while others do not, and this entails important differences in their behavior. Specifically, I will suggest that only full-fledged DPs access the specifier position of an object agreement projection.

1. Previous analyses

1.1. The paradigms

Table (1) below shows the two paradigms in question, for the verb *lát* 'see', in present tense. The table is set up according to the number and person of the subject governing agreement. With other tenses/moods, and with front vowel harmony, some of the endings are slightly different, but these alterations do not affect our discussion and results in any way.

(1)

	Subjective			Objective	
lát 'see'	SG	PL		SG	PL
1st	lát-ok	lát-unk		lát-om	lát-juk
2nd	lát-sz	lát-tok		lát-od	lát-játok
3rd	lát	lát-nak		lát-ja	lát-ják

The 'subjective' forms in the table have no correlation whatsoever with any property of any other phrase than the subject, however, these forms are used (among other cases) when the object is a 1st or 2nd person non-reflexive personal pronoun, except for the single

This paper is a more experimental version of Bartos (to appear). I wish to thank Katalin É. Kiss for encouraging me to write it, and discussing it with me in detail. I am also grateful to Ágnes Bende-Farkas, Anna Szabolcsi, and Gabriella Tóth for helpful discussions, and to Michael Brody, László Kálmán, András Komlósy, Gréte Dalmi, and Viktor Trón, for their various comments and suggestions. Much improvement in the quality of this paper is due to them, while all inadequacies and errors are mine. Finally, I wish to mention Teun Hoekstra, whose personally communicated ideas on Hungarian inspired me a lot.

case when the subject is 1st person singular and the object is 2nd person—in this case the form *lát-lak* 'I-see-you' is used. This is the sole occurrence of clear person agreement with the object.[1] As regards the 'objective' series, those forms do not show number and/or person agreement with the object, in the strict sense, either. On the one hand, though it is true that they basically stand with 3rd person objects, reflexives in any person (and reciprocals) trigger this paradigm, as well. On the other hand, it is not the case that any 3rd person object forces the objective inflection—as will be discussed in much detail below, indefinites in many cases cooccur with the subjective paradigm. Thus we can immediately conclude that any attempt to explain the distribution of the two paradigms in terms of number/person object agreement is flawed.

1.2. Definiteness agreement?

The second usual analysis of the phenomenon relies on the notion of definiteness of the object: roughly speaking, if the object is a definite NP, it goes together with 'objective' agreement on V, whereas if it is indefinite, the 'subjective' inflection is chosen, cf. (1).[2] (This leading idea is implemented, with different details, for example in Rácz & Takács (1974) [a brief reference grammar], Szamosi (1976), and, at least for 3rd person objects, in Szabolcsi (1992, 1994a), Farkas (1987).[3]

(1a) Látom / *látok a fiút.
see-1sg-ob / see-1sg-sub the boy-acc
'I see the boy.'

(1b) Látok / *látom egy fiút.
see-1sg-sub / see-1sg-ob a boy-acc
'I see a boy.'

Furthermore, intransitive verbs pattern with verbs taking an indefinite object in this respect. This last fact is in itself a weak point of this analysis, in as much as it needs to be stipulated, since it is less than obvious that if the key factor in the choice bewteen the paradigms is definiteness, then intransitive verbs should choose the 'indefinite' agreement affixes. Not having any object, they might as well go with the 'definite' agreement endings—the sole thing that could be evoked to remedy the situation is markedness, provided we rightfully regard the objective paradigm as more marked than the subjective one.[4]

---

[1] The suffix *-lak* can in fact be broken up into *-l-*, which is one variant of the marker of 2nd person, in the subjective paradigm (taking the place of *-sz* seen in Table (1) after stems ending in sibilants), followed by *-a-*, possibly analyzed as an epenthetic vowel, and the final *-k*, i.e. the 1st person subject agreement suffix (cf. the subjective endings).

[2] In the glosses all number/person agreement specifications are meant as agreement with the subject, unless explicitly indicated otherwise; 'sub' and 'ob' mark 'subjective' vs. 'objective' inflection; features (other than agreement) not overtly marked on a particular form, e.g. present tense indicative, are dropped from the glosses. Also, Hungarian displays no gender distinctions, not even on pronouns; for simplicity's sake I will use the masculine forms in the glosses and translations throughout.

[3] In the latter two, it is necessarily assumed that specific indefinites, discussed below, formally/featurally count as definites.

[4] On a markedness account see Moravcsik (1988).

There are several empirically rooted objections to the definiteness agreement hypothesis, too. Definiteness of a nominal phrase is to a large extent the function of the determiners. For instance, determiners such as *egy* 'a/one', *néhány* 'some', *öt* 'five', are called indefinite determiners, in keeping with the assumption that they render the NP they determine indefinite. Thus, as expected under the definiteness analysis, they occur with subjective agreement on the verb as objects; cf.

(2) Látok / *látom öt embert.
see-1sg-sub see-1sg-ob five man-acc
'I see five men.'

However, when the object includes a possessive construction, the verb usually appears with the objective paradigm, even though the same indefinite determiner is present (and, accordingly, the NP is still interpreted as indefinite), as in (3):

(3) Látom öt emberedet.
see-1sg-ob five man-2sgPOSS-acc
'I see five of your men.'

In fact, in such cases the verb could carry subjective endings, too, but with a different (non-specific) interpretation. This contrast will be treated below in detail.

A similar case is shown, this time with an indefinite pronoun, in (4a) vs. (4b):

(4a) Látok / *Látom valakit.
see-1sg-sub see-1sg-ob someone-acc
'I see someone.'

(4b) Látom valakidet.
see-1sg-ob someone-2sgPOSS-acc
'I see someone belonging to you.'

Once again, the inherent indefiniteness of the object does not fully determine the choice of agreement paradigm—instead, other factors need to be considered, too. (And once again, in (4b), objective inflection coud be used, but with a shift in the specificity of the object.)

Another complication with a definiteness account is caused by the determiner *minden* 'every'. Normally, *minden* triggers subjective agreement:

(5) Látunk / *látjuk minden fiút.
see-1pl-sub see-1pl-ob every boy-acc
'We see every boy.'

This situation changes, however, in certain cases. For example, similarly to the above instances, the presence of a possessive construction results in a switch to objective agreement, as in (6a). Likewise, if *minden* is preceded by the definite article[5], the objective

[5] *Minden* (and a number of other determiners) cannot be directly preceded by the definite article, unless there is some intervening material between them. Szabolcsi (1994) offers a phonological account

pattern appears, cf. (6b). That definiteness should not be a decisive factor here is illustrated by (6c), a minimally differing case, requiring subjective conjugation.

(6a) Ismerem (a te) minden titkodat.
know-1sg-ob (the you-NOM) every secret-acc
'I know your every secret.'

(6b) Elégetem / *elégetek a tőled kapott minden levelet.
burn-1sg-ob burn-1sg-sub the from-you received every letter-acc
'I burn every letter received from you.'

(6c) Elégetek / *elégetem minden tőled kapott levelet.
burn-1sg-sub burn-1sg-ob every from-you received letter-acc
'I burn every letter received from you.'

Finally, there are interesting cases, with a possessive construction lacking both an overt possessor, and an overt article, where the subjective paradigm optionally steps in, at least in certain dialects of Hungarian, including my own:

(6d) Ismerek (*a te) minden titkodat.
know-1sg-sub (the you-NOM) every secret-acc
'I know your every secret.'

Clearly, then, neither definiteness itself, nor the possessive construction (possibly seen as giving rise to definiteness), on its own, can be used as an explanation for the distribution of objective agreement.

A further problem is posed for the definiteness agreement hypothesis by the fact that 1st and 2nd person personal pronouns, when objects, occur with the subjective agreement pattern, witness (7a), as opposed to 3rd person object pronouns (7b).

(7a) Péter lát / *látja engem / téged / minket / titeket.
Peter see-3sg-sub see-3sg-ob me you(sg)-acc us you(pl)-acc
'Peter sees me / you(sg) / us / you(pl).'

(7b) Péter látja / *lát őt / őket.
Peter see-3sg-ob see-3sg-sub him them
'Peter sees him / them.'

It seems perfectly unreasonable to draw a distinction between 1st and 2nd person pronouns, on the one hand, and 3rd person ones, on the other, in terms of definiteness.[6]

for this, claiming that there is nothing inherently wrong in the cooccurrence of the two, and in fact the article is there for syntactic and semantic purposes, but a PF-filter blocks them from appearing adjacent to each other, and deletes the article in those cases, while if there is some lexical material between them, the article can stay.

[6] As Farkas (1990) notes, 1st and 2nd pereson pronouns can be pro-dropped, and since pro-drop in Hungarian is confined to definites, this is a syntactic argument, added to the obvious semantic argument, for regarding these personal pronouns as definite.

The only phenomenon that may suggest so is exactly the one in question, namely the divergence in the choice of V-agreement paradigms.

Finally, in some dialects[7] there is an interesting contrast correlating with the alternation of agreement endings, but (crucially) not involving any necessary difference in definiteness, as shown in (8a) vs. (8b):

(8a) Olvastuk Péter (öt) versét.
read-past-1pl-ob Péter(NOM) (five) poem-3sgPOSS-acc
'We have read Peter's (five) poems.'

(8b) Olvastunk Péternek (öt) versét.
read-past-1pl-sub Peter-dat (five) poem-3sgPOSS-acc
'We have read (five) poems by Peter.'

This contrast seems to be attributable to a difference in the **specificity** of the object. In the absence of anything better, we may be inclined to say at this point that the specific–non-specific distinction plays a role in the choice between the objective and the subjective paradigms.

1.3. Specificity agreement?

In the light of the problems discussed above, it is a natural move to examine the possibility that Hungarian 'object agreement' is at least partially a case of specificity agreement. More precisely, one might claim either that (i) the prime factor governing object agreement is definiteness, but under certain conditions (especially in the case of indefinite objects) specificity may intervene, or that (ii) specificity, rather than definiteness, is the key feature. Let us take a look at the previously mentioned problems once more, to see whether we are any better off with (i) or (ii).

As it happens, (2) and (4a) are immediately problematic for a 'specificity only' approach. The object phrases *öt ember* 'five men' and *valaki* 'someone' are ambiguous in this respect: they can be interpreted either specifically or non-specifically, however, they will invariably trigger subjective agreement. Moreover, the object in (3), albeit a possessive construction, is not necessarily any more specific than the one in (2), yet it tends to occur with objective agreement. A combined definiteness-and-specificity account may be more viable, as long as we can maintain that with non-possessives definiteness counts, and with indefinite possessives paradigm selection hinges on specificity. Definite possessives are obviously specific. The data in (6), however, gets us into trouble. Arguably, there is no definiteness or specificity difference between the objects of (6b) and (6c), yet the contrast in agreement patterns is perfectly clear.

It is necessary to make mention of Enç's (1991) concept of specificity, where a nominal phrase counts as specific iff its discourse referent is linked to some previously established discourse referent by a relation of inclusion, as opposed to the case of definites, where the relevant linking relation is identity. Now, it might seem promising to follow a line here building on the assumption that possessedness in fact satisfies the criteria of the inclusion

[7] In most dialects, including 'standard' Budapest speech, (8b) is not acceptable, and (8a) can have both readings; furthermore, even in the 'contrast' dialects, (8b) is merely an option to express non-specificity—(8a) will do as well for this purpose.

relation, hence the possessive constructions would immediately qualify as specific, rightfully triggering objective agreement under a specificity approach. Enç's theory is all the more attracting, because it is syntactically anchored: in Turkish, specific objects stand with a distinctive case-suffix, in opposition to non-specific ones, which always occur bare. Hungarian thus apparently parallels the situation in Turkish, the difference being that here it is verbal agreement, rather than case, morphology is the signal. However, the fact that (1b) and (2) are variably interpretable with respect to the specificity of the object without any concomitant alternation in agreement does not easily yield itself to a neat explanation in Enç's terms, besides, universal quantifiers show a striking mismatch: in Turkish they behave morphologically as specifics, and Enç actually argues that also from a semantic point of view they induce specificity. But in Hungarian, as (5) and (6) show, they clearly pattern with non-specifics. It is therefore reasonable to look for a better characterization of the Hungarian agreement choice than specificity.

The best we can apparently say is that somehow the overt definite article counts for agreement. But this is worth nothing under Szabolcsi's (1994) theory, where the definite article is always present with *minden* 'every', except at PF, thus there can be absolutely no difference there in syntax, and paradigm selection presumably takes place before the output of morphology is fed into PF.

Furthermore, the split of personal pronouns remains a problem, unless one wants to claim that there really exists some specificity difference between 3rd person pronouns and the rest.[8] Eventually, the data in (8) proves to be the only compelling motivation for seeking the solution in terms of specificity.

So what the data suggests is that although definiteness and specificity do show some correlation with the choice of object agreement, it is worth investigating other options, whereby it may turn out that this correlation is in fact an effect, rather than the cause.

At this point, before proceeding to my proposal, it seems useful to highlight the most crucial questions lacking a good answer:

(i) Why do intransitives pattern with transitives taking a 'definite' (or 'specific') object in choosing the subjective conjugation?

(ii) Why does the possessive construction trigger the objective paradigm, and why is the (6d, 8b)-type an exception to this?

(iii) Why is there a split between 3rd person and non-3rd person pronouns, in that the latter pattern with 'indefinites', requiring subjective agreement?

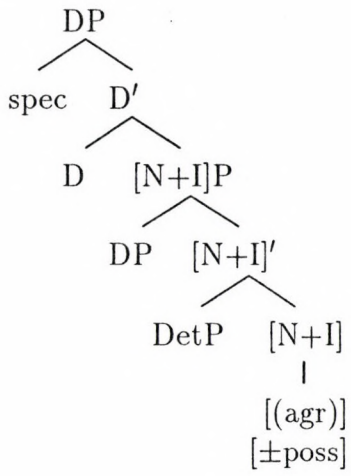
2. The proposal

2.1. A generalization

For what follows, I adopt the phrase structure attributed to nominal phrases in Hungarian as presented in Szabolcsi (1992, 1994a), shown here in (9).

[8] É. Kiss (p.c.) suggests that one might toy with the idea of taking 1st and 2nd person pronouns to be non-specific, in a discoursal sense, on the grounds that they can never be coindexed with a syntactic antecedent—the sole way of rendering an NP specific. Another suggestion (Jeffrey Goldberg, p.c.) segments the specificity hierarchy into three parts, with the 1st and 2nd person pronouns, being at the [+specific] extreme, constituting a third class, an indication of which is the fact that with a few optionally transitive verbs, like "it eszik 'eat', in the case of 3rd sg. subjects, they stand with a verb-form belonging to a 'third paradigm': the ending is different from both the 'subjective' and the 'objective' inflection, cf. "itemitem(i)"it Esz / *eszi / ?eszik engem a méreg 'eat-3sg(-*ob/?sub) me the anger.'

(9)



An important property of this analysis is the strict separation of two classes of determiners. One class comprises the definite article *a(z)* 'the', the zero indefinite article, and the demonstrative+article complex *ez/az a(z)* 'this/that–the'—their category is D^0, and they head the outmost projection of nominal phrases. In terms of distribution, they always precede nominative-marked possessors.

(10a) [$_{DP}$ a [$_{[N+I]P}$ mi barátunk]]
the we(-nom) friend-1plPOSS
'our friend'

As opposed to this group, there is another class, including simple demonstratives (e.g. *e(me), ezen* 'this', *ama, azon* 'that'), quantifiers (e.g. *minden* 'every', *kevés* 'few', *egy(ik)* 'one'), and numerals (e.g *öt* 'five'). These are full maximal projections (DetPs in Szabolcsi's term), and occupy a slot following nominative-marked possessors.

(10b) [$_{DP}$ a [$_{[N+I]P}$ mi [$_{[N+I]'}$ minden/kevés/öt barátunk]]]
the we(-nom) every/few/five friend-1plPOSS
'our every/few/five friend(s)'

They are within a maximal projection smaller than DP; I will tentatively assume with Szabolcsi that they are in [N+I]P, whose head is an [N+I] complex (where I is the possessive–agreement inflection) and whose specifier is filled by the nominative possessor.[9]

2.1.1. Non-possessives

Considering now the simple cases of 'object agreement', where no possessive construction is involved, we get a straightforward account on the following basis: whenever there is an overt D^0 in the object phrase, objective agreement is forced on the verb, and subjective agreement is the elsewhere case. Assuming a principle of projectional economy (see e.g. Grimshaw 1991, 1993), we can rephrase the situation, saying that whenever the object is a fully projected nominal phrase, i.e. a DP, it triggers objective agreement, and when it is

[9] Szabolcsi (1992, 1994a) attributes entirely different functions to these classes. She argues that D^0s are pure subordinators, not determiners in the semantic sense, while instances of DetP are determiners, and may consist merely of features like [±definite], [±specific], in association with the 'definite' article occupying D^0, hence the apparent role of the article in determining definiteness and specificity.

not a full-fledged DP, i.e. a smaller nominal projection, such as [N+I]P, it does not—the default case being subjective agreement.[10]

At this point it is clear already, why definiteness of the object nominal correlates with the paradigm selection. Either the article in D^0 is itself the source of definiteness, or (in keeping with Szabolcsi (1992, 1994a)) there are matching rules between D^0 and DetP which ensure that the definite article only occurs when the DetP specifies its containing [N+I]P as [+definite] (or at least [+specific]).

Next we should tackle indefinite, interrogative, negative, universal, and relative pronouns, which always occur with subjective agreement, unless they are placed into a possessive construction as the possessed element. ((4a) is repeated here as (11a).)

(11a) Látok / *Látom valakit.
see-1sg-sub see-1sg-ob someone-acc
'I see someone.'

(11b) Kit látsz / *látod ?
who-acc see-2sg-sub / see-2sg-ob
'Who do you see?'

(11c) Senkit nem látok / *látom.
nobody-acc not see-1sg-sub / see-1sg-ob
'I see nobody.'

(11d) Mindenkit látok / *látom.
everyone-acc see-1sg-sub / see-1sg-ob
'I see everyone.'

(11e) akit látsz / *látod
who_rel-acc see-2sg-sub see-2sg-ob
'who(m) you see'

The internal structure of these pronouns is not perfectly clear, but we may build on Cheng's (1991, p. 84 ff.) analysis, and claim that at least in (11a–d), the pronouns are Det+NP complexes, where NP is *kit*, a unit without quantificational force, and Det (*vala-*, *∅-*, *sen-*, *minden-*, respectively) is a quantifier (indefinite, wh-, negative, and universal, respectively). Det thus falls in with DetPs in Szabolcsi's DP-structure (in the case of *minden* it is even the same form), thereby all of these pronouns are just [N+I]Ps, not DPs, insofar as overt material is concerned. It is no surprise, then, that they do not trigger objective agreement.

Additional support for my hypothesis comes from incorporated objects, as illustrated in (12). (For a discussion of these, see e.g. É. Kiss (1992, 1994).)

[10] The only problem with this view is that Szabolcsi admits a null indefinite article among D^0s, one possible reason for which is that a SpecDP position (hence a D^0) is needed for allowing a possessor to leave the nominal phrase (to topicalize, for instance). Since I offer a different analysis for this phenomenon, I regard the null article as not present at all. The semantic consequences of omitting this null element, and the whole projection it would head, do not concern me here.

(12) Almát eszünk / *esszük.
apple-acc eat-1pl-sub / eat-1pl-ob
'We are eating apples. (We are apple-eating.)'

As seen in the example, these bare nominals never stand with objective agreement. Since they are just X^0s, this is what we expect.[11]

2.1.2. Possessives

Let us now turn our attention to possessives. Recall that in some of these cases there is an option whether such objects stand with subjective or objective agreement. The first-sight generalization seems to be that an overt definite article, or an overt nominative-case possessor, requires objective inflection (13a, b), while in the absence of both, that is, when the possessor is non-overt, or dative-marked and outside the object phrase, both agreement paradigms are grammatical, but with a specificity difference on the object (cf. (8a, b)).

(13a) Látom / *látok a kutyádat.
see-1sg-ob / see-1sg-sub the dog-2sgPOSS-acc
'I see your dog.'

(13b) Látom / *látok Péter kutyáját.
see-1sg-ob / see-1sg-sub Peter(-nom) dog-3sgPOSS-acc
'I see Peter's dog.'

The presence of an overt D^0 fits the scheme sketched above: it necessitates the projection of the DP-layer. Without it, it is at least possible for the nominal phrase to lack this outermost layer. More trouble is caused by the possessors. In Szabolcsi's now standard analysis (for details see e.g. her (1994)), the nominative-case possessor occupies the specifier of [N+I]P ((14a)), while its dative-case counterpart is found in the spec of DP, if it is still within the DP at all ((14b)), for it is capable of leaving the DP altogether, by way of operator-type movements (such as topicalization, focusing, left-dislocation), or scrambling ((14c)).

(14a) $[_{DP}\ [_{D}$ a] $[_{[N+I]P}$ Péter $[_{[N+I]'}$ kutyája]]]
the Peter(-nom) dog-3sgPOSS
'Peter's dog'

(14b) $[_{DP}$ Péternek$_i$ $[_{D}$ a] $[_{[N+I]P}\ t_i\ [_{[N+I]'}$ kutyája]]]
Peter-dat the dog-3sgPOSS
'Peter's dog'

(14c) $[_{CP}$ Péternek$_i$ $[_{VP}$ eltűnt $[_{DP}\ t'_i\ [_{D}$ a] $[_{[N+I]P}\ t_i\ [_{[N+I]'}$ kutyája]]]]
Peter-dat disappeared the dog-3sgPOSS
'Peter's dog disappeared'

[11] É. Kiss (1992) regards them as XPs represented solely by their heads. If so, they are probably the minimal XPs, i.e. [N+I]Ps in our case, absent any evidence to the contrary, so they pose no problem for my analysis. But they occupy the same slot as verbal prefixes do, moreover they can be considered to be fully incorporated into V, which suggests that they may turn out to be mere X^0s.

The two positions cannot normally be filled simultaneously. Though it is possible to have the full possessor phrase in the spec of DP, and a coreferential (resumptive?) pronoun in the spec of [N+I]P, it is markedly archaic, or jocular, in flavor, cf. (15):

(15) % Péternek$_i$ az ő$_i$ kutyája
Peter-dat the he(-nom) dog-3sgPOSS
'Peter's dog' (lit.: 'Peter's dog of his')

Szabolcsi, therefore, assumes that the two positions are movement-related: all possessors originate in the inner position, and can actually stay there, receiving nominative case, but they can (or in certain cases: must; see below) raise up to the outer position, which is somehow associated with a dative(-like) ending, and which can serve as an escape hatch for further movement. Also, the outer position is an operator position (which the inner one isn't).

This picture is totally incompatible with my proposal, because (i) nominals with a nominative possessor and without an overt D^0 would count as less-than-DPs, and would thus occur with subjective agreement, contrary to the facts; and (ii) dative-marked possessors would imply the presence of the DP-layer, being in need of a SpecDP, so subjective conjugation (as in (8b)) should be impossible with them. For the latter, I assume that in (8b)-type cases there is no DP projected, rather, the possessor moves directly out of the [N+I]P-internal position. Overt D^0 is never found in these cases (that would immediately trigger the objective conjugation, and yield a definite interpretation). The problem we face now is how to explain the fact that nominative possessors cannot be extracted, cf. (16).

(16a) *Péter$_i$ olvastunk [t_i versét].
Peter(-nom) read-past-1pl-sub poem-3sgPOSS-acc
'We have read poems by Peter.'

(16b) Péternek$_i$ olvastunk [t_i versét].
Peter-dat read-past-1pl-sub poem-3sgPOSS-acc
'We have read poems by Peter.'

(16c) *Péter$_i$ olvastuk [(a) [t_i versét]].
Péter(-nom) read-past-1pl-ob (the) poem-3sgPOSS-acc
'We have read Peter's poem.'

(16d) Péternek$_i$ olvastuk [t'_i (a) [t_i versét]].
Péter-dat read-past-1pl-ob (the) poem-3sgPOSS-acc
'We have read Peter's poem.'

In Szabolcsi's account this followed from the fact that the extracted possessor had to pass through SpecDP, where it picked up its dative-ending. On the other hand, her theory does not explain why the possessor has to be extracted when D^0 is a null-element (=[−specific]), i.e. why extraction is obligatory for a non-specific reading to arise (Szabolcsi 1994a, p. 227). This is evident here, since with the 'null' D^0 there is no D-projection, hence no

SpecDP, while an in-situ, nominative possessor would force the specific reading. The reason why the nominative possessors fail to move, under minimalist assumptions (Chomsky 1995), must be that they have nothing to check, neither Case, nor operator features.[12]

We now have to say something about problem (i), i.e. the obligatory 'DP-ness' of nominative-possessor phrases. It is clear that if the possessor is non-overt (i.e. pro), then all depends on the presence vs. absence of an overt D^0, as shown in (17):

(17a) Láttunk / *láttuk kutyádat.
see-past-1pl-sub / see-past-1pl-ob dog-2sgPOSS-acc
'We have seen some dog(s) belonging to you.'

(17b) Láttuk / *láttunk a kutyádat.
see-past-1pl-ob / see-past-1pl-sub the dog-2sgPOSS-acc
'We have seen your dog.'

This neatly corresponds to the DP vs. [N+I]P difference. Furthermore, if the [N+I]P-internal possessor is an overt personal pronoun, the definite article must be present, and consequently the objective agreement and the definite reading is the only option:

(18) Láttuk / *láttunk a te kutyádat.
see-past-1pl-ob / see-past-1pl-sub the you(-nom) dog-2sgPOSS-acc
'We have seen your dog.' (*'We have seen some dog(s) belonging to you.)'

This fact may serve as an indication that overt nominative possessors necessarily occur in DPs, even if in many cases there is no overt D^0, cf. (19):

(19a) Láttuk / *láttunk (a) Péter kutyáját.
see-past-1pl-ob / see-past-1pl-sub (the) Peter(-nom) dog-3sgPOSS-acc
'We have seen Peter's dog' (*'We have seen some dog(s) of Peter.')

(19b) Láttuk minden / egy / a / sok fiú kutyáját.
see-past-1pl-ob every / a / the / many boy(-nom) dog-3sgPOSS-acc
'We have seen every/a/the boy's / many boys' dog.'

In some of these cases one might argue (following Szabolcsi 1992, 1994a) that the definite article is present in syntax, and deletes at PF, obeying a rule of 'haplology', the function of which is to eliminate D–D and D–Det sequences.[13] Even for (19a) one could propose that proper names like 'Peter' inherently contain a definite article, and even this can trigger the PF deletion rule. But surely there is nothing wrong with D–Num, or D–*sok* ('the

[12] It is not inconceivable that there is some functional head (and the corresponding projection) in nominal phrases (perhaps Agr or Case) and the nominative case possessors in fact check Case in its specifier, then stay put. Alternatively, it may be the case that nominative NPs/DPs are syntactically Caseless, as proposed by Bittner & Hale (1996).

[13] Indeed, sequences like *az egy* 'the a/one', *a minden* 'the every' are very rare in Hungarian, and two subsequent definite articles are totally impossible, even if such a sequence is syntactically and semantically plausible, as in *a* [*a fiú*] *kutyája* 'the [the boy(-nom)] dog-3sg"smc poss', meaning 'the dog of the boy' (note the double occurrence of 'the' in the translation).

many') strings, so Szabolcsi's idea cannot be extended to these cases. We are certainly short of a perfect explanation here.

Yet some support to the underlying presence of a D^0 comes from the fact that for each of these cases a dative-marked possessor in SpecDP is an alternative option, followed by an overt definite article, with no meaning difference at all, which is suggestive of the presence of D^0 with the nominative-case possessors, too.

2.4. Remnants

There are a few other cases that have not been specifically mentioned up to this point, but merit some discussion. One of these is the fact that there are certain possessors that cannot appear in the [N+I]P-internal position, only in SpecDP, or outside of the nominal phrase, with a dative-ending. These include indefinite, negative, interrogative, universal and relative pronouns:

(20a) *a ki/valaki fia
the who/someone(-nom) son-3sgPOSS
'whose/someone's son'

(20b) kinek/valakinek a fia
who-dat/someone-dat the son-3sgPOSS
'whose/someone's son'

Here I follow Szabolcsi (1994) in attributing the phenomenon to the operator nature (=feature) of these elements, as well as of the positions they occupy: SpecDP, and the DP-external, clause-level specifiers.

Another interesting question is why object clauses mostly trigger objective agreement, as shown in (21):

(21a) Tudom / *tudok [(azt) hogy Péter okos.]
know-1sg-ob / know-1sg-sub (it-acc) that Peter smart(-sg)
'I know that Peter is smart.'

(21b) Pétert$_i$ akarom / *akarok [hogy megverd t_i].
Peter-acc want-1sg-ob / want-1sg-sub that beat-imp-2sg
'It is Peter that I want you to beat.'

In the detailed analysis of Hungarian embedded clauses, Kenesei (1992) proposes to treat *that*-clauses as [DP, CP] chains, where CP is theta-marked by the matrix V, while DP is in a Case-position, Case-marked by the matrix V. In (21), *azt* 'it-acc'(an expletive) and *Pétert* 'Peter-acc' represent this DP. Consequently, object agreement holds with this DP. *Az* 'it' is a DP-equivalent pronoun, and *Péter* is a proper name, i.e. a DP, inherently, so objective conjugation is expected. If, however, this position is taken up by a phrase that counts, in the sense of the discussion above, as less (or other[14]) than DP, subjective agreement is what we expect, and it is what we find:

[14] This, with the example in (22c), was pointed out to me by a reviewer.

(22a) Kit$_i$ akarsz [hogy megverjek t_i]?
who-acc want-2sg-sub that beat-imperative-1sg
'Who do you want me to beat?'

(22b) Öt fiút$_i$ akarok [hogy megverj t_i].
five boy-acc want-1sg-sub that beat-imperative-2sg
'I want you to beat FIVE BOYS.'

(22c) Hallottál olyat [hogy egy elsős okos legyen] ?
hear-past-2sg-sub such-acc that a first_grader smart be-imper-3sg
'Have you ever heard such a thing that a first-grader should be smart?'

To sum up briefly, these cases do not constitute counter-evidence; their behavior is in full compliance with our theory, once we have a correct analysis for them.

3. A minimalist analysis

3.1. DPs, Case, and object agreement

In this section I turn my attention to the technicalities of implementing my proposal in a minimalist framework, the basics of which are found in Chomsky (1995). In keeping with the currently standard assumptions about the functional structure of clauses, I posit an object agreement functional head and projection: Agr$_O$^{0}, and Agr$_O$P, and claim that Agr$_O$ is the locus of checking the object agreement features on the verb, directly related to the 'subjective' vs. 'objective' inflectional morphology. One way to move towards an account of the phenomena discussed above would be to link Case-licensing with object agreement—as is standard in most versions of Minimalism.[15] We could argue that certain object phrases, which are not DPs, just NPs or [N+I]Ps, do not check features at SpecAgr$_O$, thus do not license objective agreement on V. In other words, they are Case-theoretically invisible to the verbal heads, unlike full DP objects, so the verbs theta-marking them will behave as intransitives from a Case-theoretic point of view. This immediately provides a simple account of why verbs taking 'indefinite' objects pattern with true (theta–)intransitives, as far as subjective vs. objective agreement is concerned. Also, if the raising of XPs to agreement- and/or Case-checking positions is driven by the connection between the attracting features of functional heads, and the D-features of the raised phrases, then it is obvious that non-DPs will not get attracted to these positions. This line of argument is pursued in Bartos (to appear). Here, however, I will try to explore another direction, where object agreement is separated entirely from Case phenomena. A motivation for this comes from Szabolcsi (1996), who assumes, on the basis of examples like (23), that indeed Case is licensed in clauses lacking an Agr$_O$ projection, as well.

(23a) El akarok kapni egy fiút.
away want-1sg-sub catch-INF a boy-acc
'I want to catch a boy.'

[15] A notable exception is the last chapter of Chomsky (1995), where the idea of agreement heads and projections is completely abandoned, and Case is associated with T, for nominative, and an outer 'light verb' of a VP-shell, for accusative.

(23b) El akarom kapni a fiút.
away want-1sg-ob catch-INF the boy-acc
'I want to catch the boy.'

Obviously, in both examples the object of the embedded, infinitival clause bears accusative ending, but the infinitive is incapable of displaying object agreement. On the other hand, the matrix verb does show object agreement variation, depending on the properties of the embedded-clause object. In Szabolcsi's terms, the object triggers agreement at the closest existing Agr_O—in this case, in the matrix clause domain. (Note that Szabolcsi's paper argues against positing a real complex predicate, consisiting of the matrix V and the infinitive, in a monoclausal structure.) The object nominal is not the argument of the matrix V, in any way, so its Case-marking must be linked to the embedded clause. That this is at least possible is confirmed by examples like (24), where the matrix predicate is thematically and Case-theoretically clearly unrelated to the embedded object in every conceivable way:

(24) El lehet kapni a fiút.
away is-possible catch-INF the boy-acc
'It is possible to catch the boy.'

Specifically, Szabolcsi suggests that apart from (and below) AgrPs, there are CasePs in the Hungarian clause structure, whose role is to license Case on nominals. Clearly, then, there are as many of these CasePs as Case-bearing phrases in the clause, furthermore, their order is arbitrary (cf. the variable order of constituents in the postverbal domain).

Splitting Case from object agreement has the advantage that uniform case morphology on all object nominals is easily captured, by assuming that accusative morphology is the spellout of $Case_{acc}$ licensed in the appropriate CaseP specifier. If Case is an N-feature, all N-projections get attracted, whether with our without a DP-layer. Agr_O, however, attracts D-features, which less-than-DPs lack, so they will not move on from SpecCaseP to $SpecAgr_O$, unlike DP objects. The nature of objective agreement is then simply the presence of an object in the checking domain of Agr_O—if this checking takes place, Agr_O is spelled out as objective inflection on V, while the absence of object agreement checking results in the appearance of subjective conjugation.

The hypothesis just sketched helps explaining an interesting paradigm, shown in connected to (22a), repeated here as (25a):

(25a) Kit akarsz [hogy megverjek] ?
who-acc want-2sg-sub that beat-IMP-1sg-sub
'Who do you want me to beat?'

(25b) Kit akarod [hogy megverjek] ?
who-acc want-2sg-ob that beat-IMP-1sg-sub
(=25a)

(25c) Kit akarsz [hogy megverjem] ?
who-acc want-2sg-sub that beat-IMP-1sg-ob
(=25a)

(25d) *Kit akarod [hogy megverjem] ?
who-acc want-2sg-ob that beat-IMP-1sg-ob
(=25a)

(25a) is the most natural one; (25b, c) are slightly less acceptable, though not really bad, while (25d) is sharply out. The wh-word *kit* is related to the embedded object position in some way. At the same time, as was seen above, it can interfere with matrix object agreement: although the object clause of the matrix V would normally trigger subjective agreement there, the wh-word, serving as the head of the chain of the embedded clause, overrides the norm, and since *kit*, an interrogative pronoun, is assumed to be a less-than-DP, subjective conjugation steps in. In particular, then, the derivation of (25a) includes the wh-movement of *kit* from the embedded object position to the matrix SpecFocusP. *Kit* is generated as the object of the embedded V, then it has to have its accusative Case licensed, either in the embedded, or in the matrix, CaseP. The example in (26) suggests that this checking is likely to occur in the matrix clause.

(26) Kit_i szeretnél [ha t_i eljönne] ?
who-acc would-like-2sg-sub if would-come-3sg
'Who would you like to come?' (lit.: 'Who would you like if came?')

The wh-phrase is generated as the embedded *subject*, of an intransitive predicate, so it cannot get its accusative case licensed there, only in the matrix clause. This implies an A-movement of the wh-phrase across the embedded clause boundary, to a matrix CaseP—an option then available for (25a), too. And since CaseP and AgrP are separate projections, a nominal phrase can check Case without influencing agreement.

In comparison, in (25b) the wh-phrase presumably checks Case in its own clause, and subsequently Ā-moves to the matrix focus position, witness the parallel (27):

(27) *Kit_i szeretnéd ha t_i eljönne?
who-acc would-like-2sg-ob if would-come-3sg
(=26)

In (25b), the expletive (*azt*) associated with the embedded clause (cf. Kenesei (1992)) is inserted into a matrix A-position, Spec of CaseP, in all likelihood, to move on to SpecAgrP and trigger the objective conjugation afterwards. Thus it excludes the wh-phrase from the matrix CaseP, whereas *kit* excludes (or deletes?) the expletive from SpecFocusP. Turning now back to (25a), we can conclude that the wh-phrase blocks the very appearance of the expletive, by occupying the matrix CaseP where it could otherwise be inserted, therefore objective agreement cannot arise in the absence of the expletive DP.

In the case of (25c, d), the objective conjugation of the embedded V is suggestive of the presence of a pro in the spec of Agr_O there, which must also have passed through the spec of the CaseP, to get its case licensed. For obvious reasons, this pro must be coindexed with *kit*, which in turn has to originate somewhere in the matrix clause, possibly as the [+wh] counterpart of the expletive *azt*. It will not be able to trigger objective agreement on the matrix V, though, while precluding the insertion of the other expletive. This is why (25c) is good, while (25d) is bad.

Another instance where the separation of CaseP and AgrP yields an easy explanation is shown in (28):

(28a) Péter van itt a legtöbbet.
Peter(-nom) be-3sg here the most-acc
'Peter is here most frequently.'

(28b) Péter hatalmasat nőtt tavaly óta.
Peter(-nom) enormous-acc grow-past-3sg last_year since
'Peter has grown enormously since last year.'

The accusative-marked phrases in these examples are not proper objects, and the verbin (28a) does not even have objective conjugation, yet the degree adverbials bear case-suffixes as 'quasi-objects'. This shows that it is not unique for the non-DP proper objects to display accusative case-endings without entering into inflectional agreement, even if they are formally DPs, as in (28a).

3.2. Remaining problems

I conclude this paper by pointing out two problem areas, where further research is necessary. One concerns the DP vs. less-than-DP distinction of nominal phrases. This distinction proved to be useful in giving an account for object agreement phenomena, but it brings its own difficulties. For one thing, if these two types are consistently distinguished, then we have to say something about why they behave identically in certain respects. One such case was brought up by M. Brody (p.c.): Although syntactic passivization has a somewhat marked (non-standard) status in Hungarian, it certainly exists, very productively, and treats my object DPs and non-DPs identically. i.e. both are potential undergoers. If Hungarian passivization is a Case-driven phenomenon, then the analysis proposed seems to provide ways to cater for it, since in this respect the DPs and the non-DPs are still alike, targeting specs of CasePs. The fact that when they are subjects, these two types do not display any divergence on the surface, is more of a worry. This leads us to the question of subject agreement, and its formalization, in Hungarian—an issue too big to be dealt with cursorily here.

Secondly, I have not offered any explanation for the fact, discussed in the first part of the paper, that 1st and 2nd person object pronouns do not stand with objective agreement, unlike 3rd person ones, which is contrary to expectations, on the assumption that they are all DP-equivalents. Note, though, that this case was equally problematic for analyses relying on definiteness, specificity, or even person/number object agreement. Farkas (1987, 1990), in fact, outlines an analysis for them in terms of feature structures, splitting apart 1st and 2nd person nominal phrases from 3rd person ones (including 3rd person pronouns) by the feature [participant].[16] Objective conjugation is triggered by a [definiteness] feature on the object, which is induced differently by the [participant] feature (for 1st, 2nd person), and by other features, like possessedness, or determiner features (affecting 3rd person nominals), so that at the point of paradigm selection 1st and 2nd person personal pronouns are not (yet) marked featurally as [definite], while at the level of semantic interpretation they (already) are. What this analysis fails to satisfactorily explain, though, is why the [participant] feature should involve this particular behavior; it is simply attributed to the "inherent definiteness" of the 1st and 2nd person personal pronouns.

To cope with the problem, I have two directions in mind, for subsequent work, to find out which (if either) is correct. One of them is to examine the categorial status of

[16] Number, i.e. plurality, is irrelevant to the issue.

1st and 2nd person pronouns: if some evidence can be found that they are less-than-DPs, then they fit into the scheme without further stipulation. The other possible path would be to relate the present facts to an 'ergative-like' split in the behavior of pronouns. The situation is further complicated by the fact that reflexive pronouns, albeit they can potentially bear 1st or 2nd person features, always occur with objective agreement, so the correct analysis cannot merely attribute the paradigm choice to person features.

Eventually, there are sporadic data that have not been treated at all. One of these is the single *látlak* 'I-see-you' form, i.e the sole case where there is person agreement with the object, besides the number and person agreement with the subject. What's more, it occurs with a 2nd person pronominal object, something that goes with subjective agreement if the subject is anything else than 1sg, by virtue of which this lonely offender should be grouped with subjective agreement, the defining property of which is the lack of checking at Agr_O! Once again, for this question to be settled, we need to examine subject agreement in detail.

Another problem area is the non-uniform behavior of *azt*. In the vast majority of cases, it stands with objective conjugation, whether it is an expletive, or a pronominal argument. There are exceptional cases, though, where both objective and subjective agreement is possible, but with a clear difference in interpretation, cf. (29):

(29a) Azt szeretnék enni, amit te.
it-acc would-like-1sg-sub eat-INF what(REL)-acc you(-nom)
'I would like to eat the same (sort of) thing as you.'

(29b) Azt szeretném enni, amit te.
it-acc would-like-1sg-ob eat-INF what(REL)-acc you(-nom)
'I would like to eat the (very) same thing/one as you.'

Yet another unexplained case is illustrated in (30):

(30) Péter jobbnak látta PRO elmenni.
Peter better-dat saw-3sg-ob away-go-INF
'Peter found it better/preferred to leave.'

The embedded infinitival clause occurs with objective agreement on this particular matrix predicate, although the usual case is for the matrix V to display subjective conjugation with infinitival complements, unless a DP object of the infinitive forces objective inflection, as we have seen.

4. Summary

I have discussed the nature of the choice in agreement inflection paradigms in Hungarian, in dependence of properties of object phrases. I have shown that previous accounts, in terms of number/person object agreement, definiteness, and specificity, are insatisfactory in some respects, and, in the case of the latter two, they are on the wrong track, in as much as correlations in these features are the result, rather than the motif, of the selection of agreement paradigms. I set up a distinction between nominals having and lacking a DP layer, and took this to be the key factor, which, through checking at an object agreement functional projection, determines the paradigm choice.

As a coda, let me reflect on the questions set up in *1.3.* I consider it one of the important gains of the proposed system, that the identical behavior of verbs without an object, and ones with an 'indefinite' object falls out trivially. I have had partial success in answering the question about possessive constructions: they take the objective conjugation, regardless of (in)definiteness, whenever they contain a D, that is, whenever they are indisputably DPs. When they are not, the possessor hangs loosely around, with a dative suffix. Finally, no satisfactory account has been found for non-3rd person pronouns, only some paths towards the solution have been sketched.

References

Bartos H.to appear. On 'Subjective' and 'Objective' Agreement in Hungarian. *Acta Linguistica Hungarica.*

Bittner, M. & K. Hale.1996. The Structural Determination of Case and Agreement. *Linguistic Inquiry* **27**, pp. 1–68.

Cheng, L.L.Sh.1991. On the Typology of Wh-questions. Doctoral dissertation, MIT, Cambridge, MA.

Chomsky, N.1995. *The Minimalist Program.* MIT Press, Cambridge, MA.

Enç, M.1991. The Semantics of Specificity. *Linguistic Inquiry* **22**, pp. 1–25

É. Kiss K.1992. Az egyszerű mondat szerkezete [The structure of simple sentences]. In: Kiefer F. (ed.), *Strukturális magyar nyelvtan 1., Mondattan [A structural grammar of Hungarian, vol.1., Syntax].* Akadémiai Kiadó, Budapest.

É. Kiss K.1994. The structure of simple sentences. In: É. Kiss K. (ed.), *The Syntactic Structure of Hungarian* Syntax and Semantics, Vol. 27. Academic Press

Farkas, D.F.1987. DO pro in Hungarian. In: Kenesi I. (ed.), *Approaches to Hungarian* vol.2., JATE, Szeged

Farkas, D.F.1990. Two Cases of Underspecification in Morphology. *Linguistic Inquiry* **21**, pp. 539–550

Grimshaw, J.1991. Extended projection. ms., Brandeis Univ., Waltham, MA

Grimshaw, J.1993. Minimal Projection, Heads, and Optimality. ms., Rutgers Univ.

Kenesei I.1992. Az alárendelt mondatok szerkezete [The structure of subordinate clauses]. In: Kiefer F. (ed.), *Strukturális magyar nyelvtan 1., Mondattan [A structural grammar of Hungarian, vol.1., Syntax].* Akadémiai Kiadó, Budapest.

Moravcsik, E.A.1988. Agreement and Markedness. In: Barlow, M. and Ch. Ferguson (eds.), *Agreement in Natural Language*, CSLI, Stanford

Rácz E. & E. Takács1974. *Kis magyar nyelvtan [A concise Hungarian grammar].* Gon-

dolat, Budapest

Szamosi, M.1976. On a Surface Structure Constraint in Hungarian. In: J.D. McCawley (ed.), *Syntax and Semantics 7: 'Notes from the Linguistic Underground'*, Academic Press, N.Y., London

Szabolcsi A.1992. Subordination: Articles and Complementizers. In: Kenesei I. and Cs. Pléh (eds.), *Approaches to Hungarian vol. 4: The Structure of Hungarian*, JATE, Szeged, 123–137.

Szabolcsi A.1994. The Noun Phrase. In: É. Kiss K. (ed.), *The Syntactic Structure of Hungarian* Syntax and Semantics, Vol. 27. Academic Press

Szabolcsi A.1996. Verb and Prefix Movement in Hungarian. Paper presented at the Research Inst. of Linguistics, HAS, Budapest; August 1996

Research Institute for Linguistics,
Hungarian Academy of Sciences,
and
Theoretical Linguistics Program,
Eötvös Loránd University
H–1014 Budapest, Színház u. 5–9

bartos@nytud.hu

THE NEGATOR-PROJECTION

By

ISTVÁN BUJDOSÓ

Abstract

In this paper I try to show the difficulties of the representation of negation. We will see that both the representation of the negator in the deep structure under VP and the conception to position the negator in the FP-adjunct position lead to contradictions. I am going to show that negator-projection is able to remove at least the technical problems. I will mention that it is not sure that precendence is the only order-principle in case of operators and their scope.

And last but not least I would like to call attention to the fact that the quetion of scope should be examined in a wider sense of the word. I expect from this act to make one more step toward a model which is able to show deeper connections between syntax and semantics.

0. Introduction

Authors who work in the framework of the GB-Theory (É. Kiss (1992, 1995), Bródy (1991)) and created successful models discuss negation rather incidentally. The aim of this paper is to find the position of the negator in Brody's model (1991), to interpret its scope and to show the connections between surface structure and logical form. Searching for the position of the negator, we have two possibilities:

1. the negator is adjucted to the negated component already under VP
2. the negated component must be moved into the scope of the negator.

In this paper I am going to argue against both of these conceptions and suppose a third one which seems to provide us with a simpler solution concerning the position of the negator and which, accommodating to the theoretical background of GB-Theory, will describe the problematic area.

1. Negator in the deep sturcture

It is a wide-spread view that the negator must be adjoined to the component it should refer to. Taking this conception, let us examine where we can position the negator in case of phrase-negation:

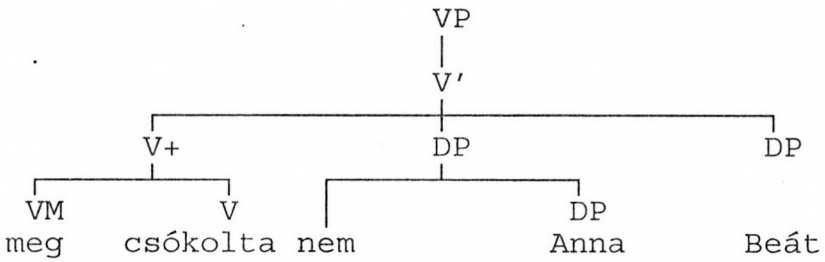
1.1. Phrase-negation

(1)	Nem	Anna	csókolta	meg	Beát.
	not	Ann-NOM	kiss-PAST	PREV	Betarice-ACC

It was not Ann who has kissed Beatrice.

The deep structure of the sentence is as follows:

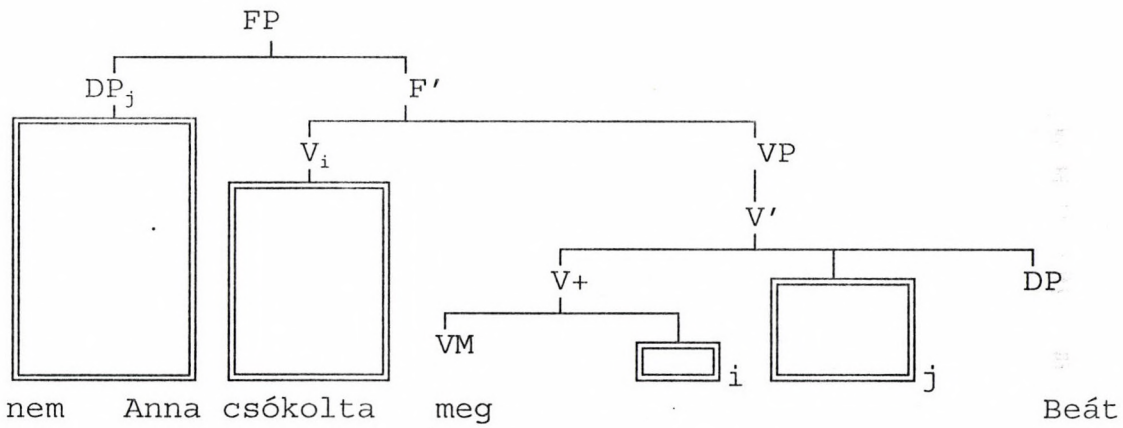
(2)



There are not many technical objections which could arise against this conception, because the generation of the sentence seems to be rather simple.

As the negated DP must be moved into the focus, the verb comes into F, and then the DP containing the negator moves into SPEC-F.

(3)



However the following grammatical problems could arise:

Is it permitted to position the negator as an operator under V' in the deep structure? The operator comes doubtless in the argument-position, which should be avoided, because the operator must bind the arguments both in logical and in grammatical sense of the word.

1.2. Sentence-negation

É. Kiss solves the problem of the sentence-negation in her 1992 model with a negator adjoined to the V'.

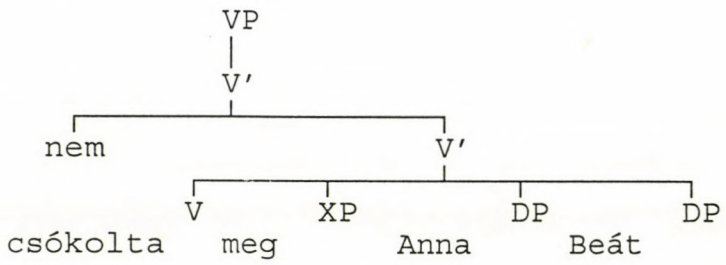
(4)

Anna	nem	csókolta	meg	Beát.
Ann-NOM	not	kiss-PAST	PREV	Beatrice-ACC

Ann has not kissed Beatrice.

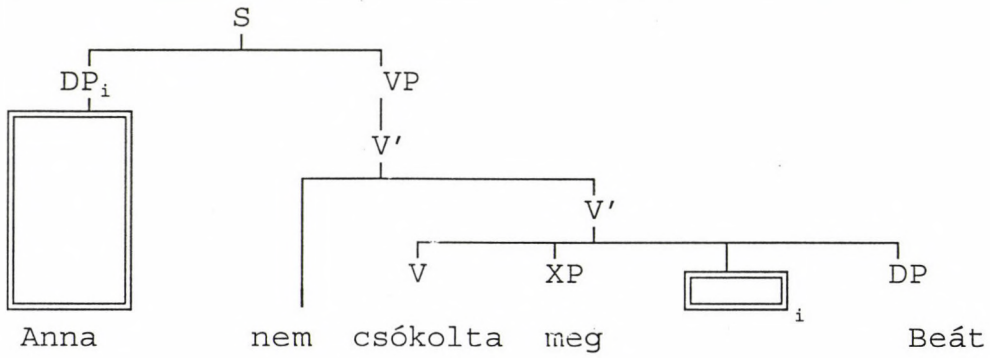
The deep structure of (4) is like:

(5)



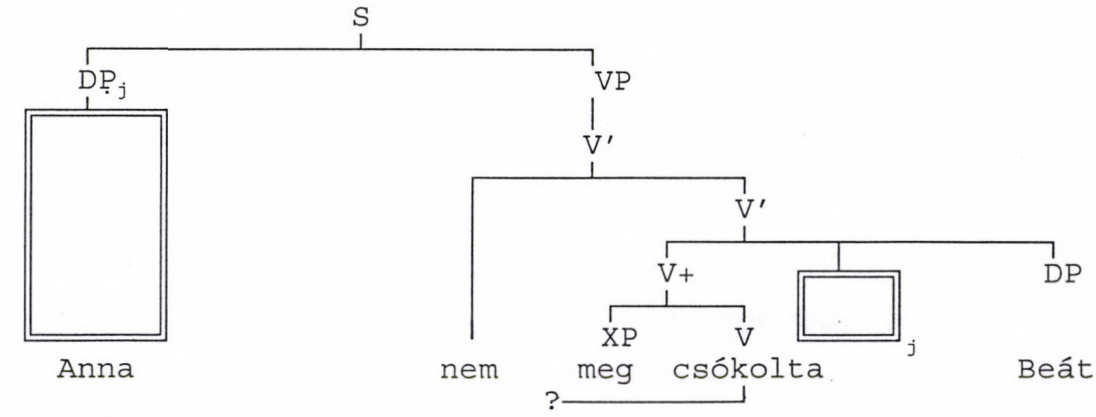
With the topicalization of the adequate DP the desired surface structure could be created:

(6)

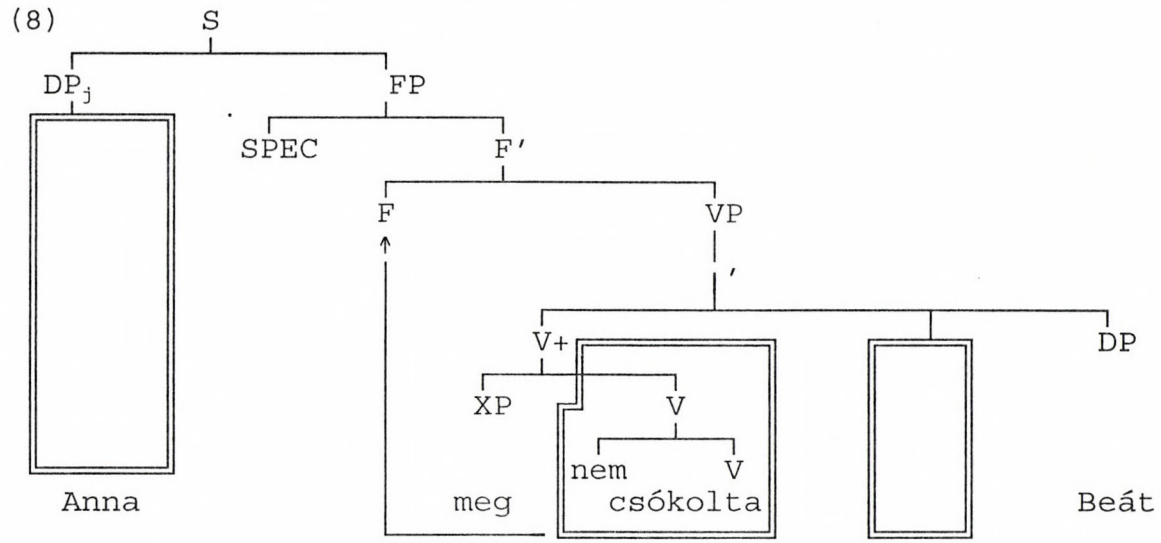


This conception cannot be held in Brody's model, because in that model the verbal prefix precedes the verb already in the deep structure. To get the desired surface structure, the verb should be moved between the negator and the verbal prefix:

(7)



But this is not possible. Therefore we must find another position for the negator. Even if the negator were adjoined directly to the verbal head, it would be difficult to find a position the verb could land on. The temptation is strong to move such verbs into the head of the focus-projection. For there cannot be found another head-position in the structure, it seems to be the only one solution:



1.3. Counter-arguments, misgivings

1.3.1. Focus in neutral sentences?

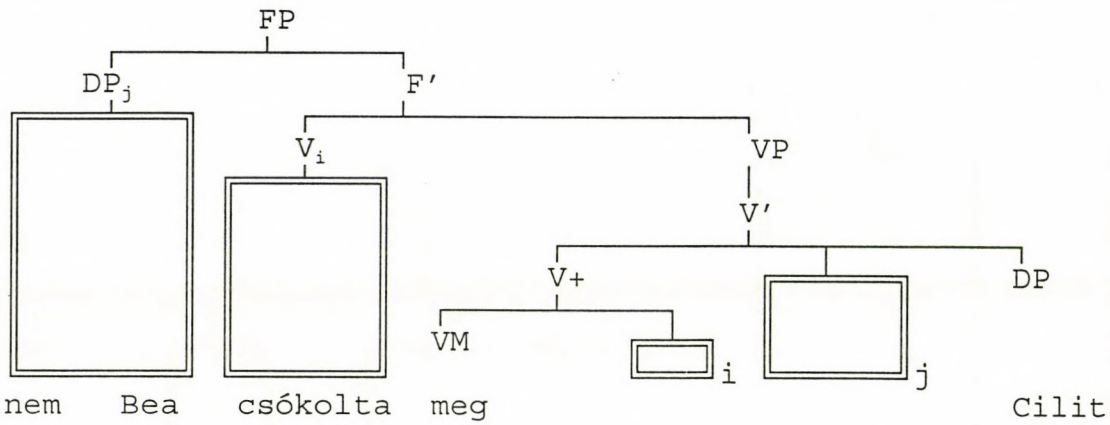
The following problem arises in (8): focus-projection can be applied only in sentences which contain a focused component, i.e. in which a focus-operator can be shown.
For the comparison let us examine some sentences in the following Montague-type world:

The model M is as usual a pair $<D,F>$, where D is a non-empty set of individuals, F an interpretation function, having as its domain the individual constants and predicates. If $\alpha \in Con_{ind}$, then $F(\alpha) \in D$; if $\alpha \in Con^k_{pred}$ then $F(\alpha) \subseteq D^k$. An assignment *g* is a function assigning an individual to each variable: if $x \in Var_{ind}$ then $g(x) \in D$.

D_e := *{Ann, Beatrice, Cecily}*
$D_t = \{0, 1\}$
$CON_e = \{a, b, c\}$
$VAR_e = \{x, y, z\}$
$CON_{<e, <e,t>>} = \{KISSED, OFFENDED\}$
$VAR_{<e, <e,t>>} = \{P\}$
$VAR_t = \{\alpha, 6, \gamma\}$
D = {Anna, Bea, Cili}
F(a) = Ann, F(b) = Beatrice, F(c) = Cecily
F(KISSED) = $\{<$Ann, Cecily$>\}$
F(OFFENDED) = $\{<$Ann, Beatrice$>\}$

In this world the following sentence is true:

(10) Nem Bea csókolta meg Cilit.
not Beatrice-NOM kiss-PAST PREV Cecily-ACC
It was not Beatrice who has kissed Cecily.



In this case it is legal to apply the focus-operator, for there is really a banishment. The focused component banishes those elements from the set of the potential "Cecily-kissers" which are not identical with itself (in this case it is Beatrice who is banished, because only this element cannot be identified as not-Beatrice). From (10) it follows (11):

(11) ANNA csókolta meg Cilit.
ANN has kissed Cecily.[1]

In (9), however, we cannot find any banishment-operator. For from the fact that "Ann has not kissed Beatrice" it does not follow that she has done something else with her. Therefore (9) is a neutral sentence, it means it is not legal to apply the focus-projection in case of the sentece-negation.

And because this was the only possibility we have found that there is no legal technical way to represent sentece-negation in Brody's model if we try to position the negator under VP in the deep structure.

1.3.2. Negative meaning

Besides the technical difficulties there are some grammar-theoretical misgivings arising against the appearance of the negator under VP, which affect phrase-negation too.

In sentences containing a negator, the negative meaning appears. This negative meaning can be defined as follows:

(15) $\exists x.\exists P(\neg Px)$ where P must not contain a negator.[2]

---

[1]Apart from the possibility that Cecily could have kissed herself.

[2] The reservation for P is necessary, because whitout this (15) would become trivial:

(16) Mirmur is a cat.
CAT(m)

From (16) follows that there exists an individual x and a predicate P, for which P is not true for x:
if x=m and P=$\neg$CAT, then $\exists x.\exists P(\neg Px)$, because $\neg\neg$CAT(m) is true.

It means: if (15) follows from the truth of a sentence, then it contains the negative meaning.

For example:

(17) Ann has not kissed Beatrice.

If we interpret x resp. P as $g(x)=a$, $P=\lambda x.KISSED(b)(x)$, then the presence of the negative meaning is already shown. For there exists an x (=a) and there exists a P ($\lambda x.KISSED(b)(x)$), for which it is true: $\neg P(x)$.
Notice that the negative meaning can be found in case of the phrase negation too:

(18) It was not Ann who has kissed Beatrice. $g(x)=a$; $P=\lambda x.KISSED(b)(x)$

The formula $\neg P(x)=\neg KISSED(b)(a)$ is still true.
Note that the negative meaning is a more abstract notion. It appears on a higher semantic level than concrete negation. It is not all-important whether we negate the whole sentence, or just a phrase (and which phrase), the negative meaning will always be true in the form $\exists x.\exists P(\neg P(x))$, if there is a negator in the sentence.
But not only in this case! The negative meaning also appears in sentences containing a focused component:

(19) ANN has kissed Cecily.

Here we can find an interpretation for x an for P in which the negative meaning is true: $g(x)=b$, $P=\lambda x.KISSED(c)(x)$, because the formula KISSED(c)(b) is furthermore not true, i.e. $\neg KISSED(c)(b)$. It means, we have found an interpretation which satisfies the formula of the negative meaning: $\exists x.\exists P(\neg P(x))$.

Returning to the syntax, we cannot find any trace of the negative meaning in the deep structure in case of the focused elements apart from the presence of the focus-projection. It would not be consequent to show the presence of the negative meaning under VP in some cases and not to do so in other cases. As it is not shown which component will be focused, it should not be shown which component will be negated either. Furthermore, note the following sentences:

(20a) Anna megcsókolta Beát.
Ann has kissed Beatrice.
(20b) Anna nem csókolta meg Beát.
Ann has not kissed Beatrice.
(20c) Nem Anna csókolta meg Beát.
It was not Ann, who has kissed Beatrice.
(20d) Anna nem Beát csókolta meg.
It was not Beatrice, whom Ann has kissed.

If (20a) is true, then the other three sentences are not true. It is not all-important what is negated. The truth value of the sentences containing a negator cannot be the same as the truth value of the sentence that does not contain any negators. It means that there must be a formula which is true for all the three sentences. This is the negated formula of the first sentence.

If the formula for (20a) is as follows:

(21) KISSED(a,b)

then the truth value of (20b-d) corresponds to the truth value of (22):

(22) ¬KISSED(a,b)

The meaning of (22) can be shown in the structure of the sentence, if the negator is positioned over the VP.

The logical form is also likely to have a deep and a surface structure. The deep structure of the logical form of the sentences (20b-d) is the same, and it corresponds to (22). That is why the aspiration for creating a common deep structure for theese sentences seems to be acceptable.

After all, the only way for us is to try to position the negator over the VP.

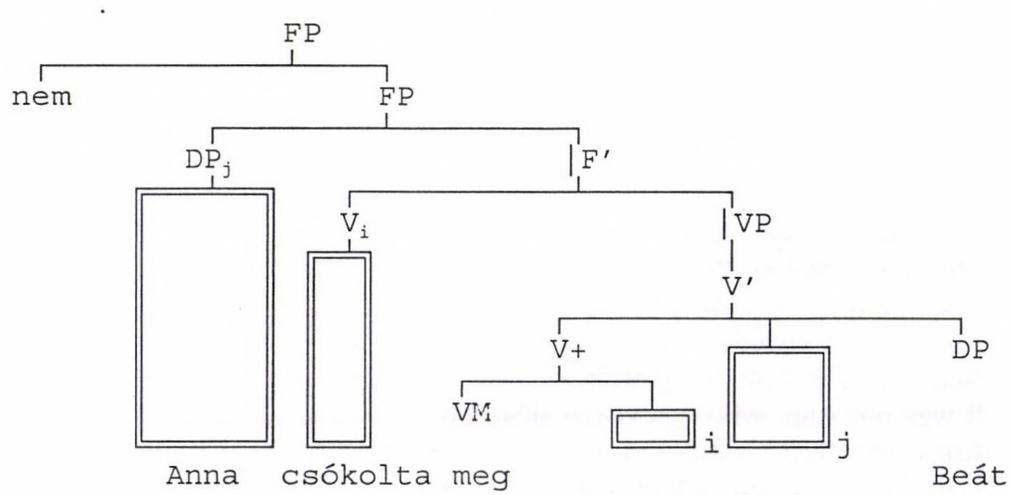
2. Negator adjoined to the FP

2.1. Phrase-negation

(1) Nem Anna csókolta meg Beát.
It was not Ann who has kissed Beatrice.

The surface structure of (1) is easy to deduce with the negator adjoined to the FP:

(23)



The solution in (23) is more advantageous than that in (2) resp. (3), because it has a more flexible relation to the phrase-negation, i.e. we do not have to decide already in the deep structure which component to negate. Therefore the deep structures of sentences which contain a phrase-negation are the same. It reflects the idea we mentioned at the end of the previous section (20a vs. 20b-d).

2.2. Sentence-negation

In case of sentence-negation we are confronted with the same problem that we had in the first section: In case of the negation of a neutral sentence, the model gets into contradiction with itself, for we cannot solve the problem without the focus-projection, which is not a legitimate solution as we have seen.

Furthermore, the appearance of a focused component and sentence-negation together cannot be solved with this configuration:

(24) ANNA nem csókolta meg Beát.
It was Ann, who has not kissed Beatrice.

because the focused component would be positioned between the negator and the verb.

2.3. Counter-arguments, misgivings

However, the following grammar-theoretical question arises in case of phrase-negation itself, which seems to be very simple:

What is the scope of the negator?

It must be admitted that the solution in (3) reflects the semantics of (1) better then the one in (23).

For the scope of the negator in (23) is as follows:

(25) nem [Anna csókolta meg Beát]
not [It was Ann who has kissed Beatrice]

According to (23) (1) refuses the following:

(26) Anna csókolta meg Beát.
It was Ann, who has kissed Beatrice.

(26) means that only Ann is an element of the set of Beatrice-kissers, i.e. the other entities are banished from this set.

Examine the logical formula for (26):

(27) $\forall x(KISSED(x,b) \leftrightarrow (x=a)) \;\&\; \exists x(KISSED(x,b)) =$
$\forall x(KISSED(x,b) \rightarrow (x=a) \;\&\; (x=a) \rightarrow KISSED(x,b)) \;\&\; \exists x(KISSED(x,b))$

The negation of this is:

(28) $\exists x(\neg(KISSED(x,b) \rightarrow (x=a))) \vee \exists x(\neg(x=a) \rightarrow KISSED(x,b)) \vee \neg\exists x(KISSED(x,b)) =$
$= \exists x((KISSED(x,b) \;\&\; \neg(x=a)) \vee ((x=a) \;\&\; \neg KISSED(x,b))) \vee \neg\exists x(KISSED(x,b))$

However, we know that somebody has kissed Beatrice, so the last member of the disjunction is false, i. e. it can be left out. The negation of (26) is therefore:

(29) $\exists x((KISSED(x,b) \& \neg(x=a)) \vee ((x=a) \& \neg KISSED(x,b))) =$
$= \quad \exists x((KISSED(x,b) \& \neg(x=a)) \vee \neg KISSED(a,b))$

It means that (23) can be interpreted in two ways:

1) There is somebody besides Ann who has kissed Beatrice,
2) or Ann has not kissed Beatrice.

According to (23) (1) is also true, if Ann and Cecily have kissed Beatrice together, or if nobody has kissed Beatrice.
If we accept (3) as the surface structure of (1), then we get the following formula:
The scope of the negator in (3):

(30) nem [Anna] csókolta meg Beát
not [Ann] has kissed Beatrice

(31) $\forall x((x=a) \rightarrow \neg KISSED(x,b)) \& \exists x(KISSED(x,b)) =$
$\forall x(KISSED(x,b) \rightarrow \neg(x=a)) \& \exists x(KISSED(x,b)) =$
$= \quad \exists x(KISSED(x,b) \& \neg(x=a))$

There is somebody who has kissed Beatrice, but he/she is not Ann.
Because (31) stands closer to our intuition than (29), we are in a fairly difficult situation. We have seen that it would be better if we tried to position the negator somewhere outside the VP. But if we do it, the problem of sentence-negation is still there, and the changed scope of the negator makes an interpretation possible which is in contradiction to our semantical expectations. And because there are no more technical possibilities in the model to solve this problem, there are two ways for us:

1) We have to change our intuition.
2) We have to change the model.

Because our intuition can not be changed, and I think the model must not be more important than the thing it models, I choose the second way.

3. The negator-projection

In section 1.3.2. we could observe in connection with the negative meaning the similarities in the behavior of the focus-operator and the negator. By reason of this it does not seem to be very constrained to suppose the existence of a negator-projection - similar to focus-projection.

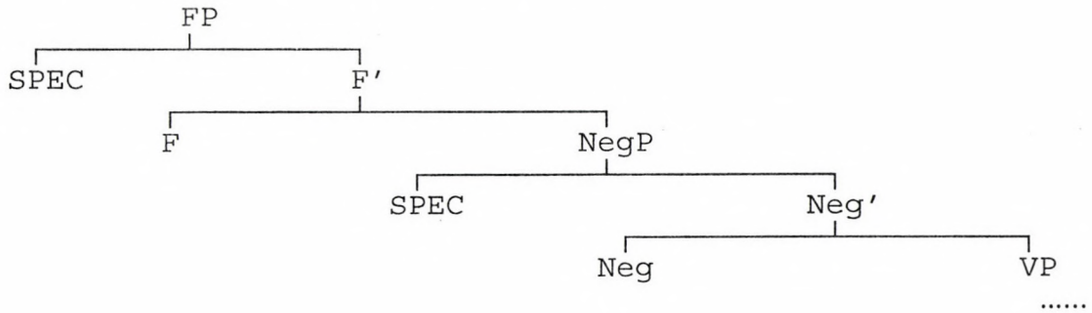
3.1. The position of the negator-projection

The negator-projection must be positioned between the FP and the VP, because - as already mentioned - sentence-negation can also be made if there is a focused component in the sentence, and in this case the focused component precedes the negated verb:

(24) ANNA nem csókolta meg Beát.
It was Ann, who has not kissed Beatrice.

Let us see how the model looks like with this addition:

(32)



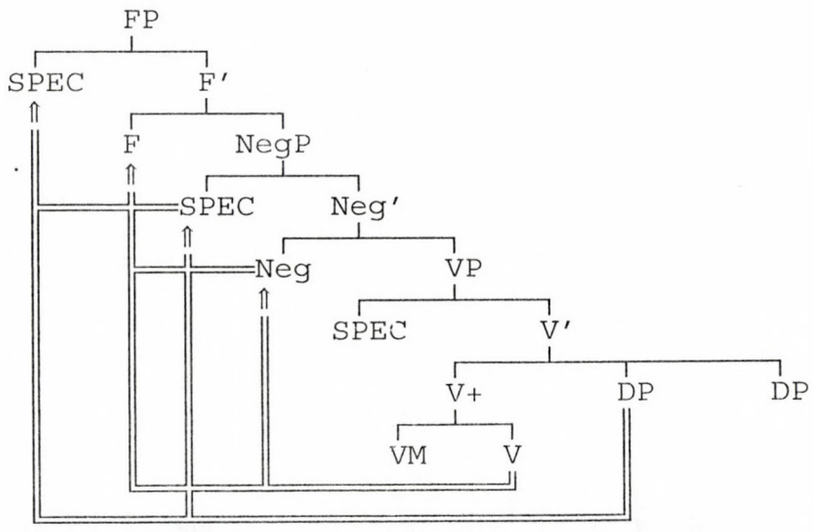
3.2. Internal structure and functioning of the negator-projection

Similar to focus-projection the head of the negator-projection also distributes a marker which could be called 'negative-marker'. The head *Neg* is able to distribute this negative-marker to the DP in its specifier position and/or to the verb which is moved into it. The effect of this negative-marker is found in the phonological form; a component X which has the negative-marker will be realized like "nem-X" (not-X).

DP	=	Anna	DP^{-}	=	'nem Anna
V	=	csókolta	V^{-}	=	'nem csókolta

With respect to the transformation-possibilities consider the following figure:

(33)

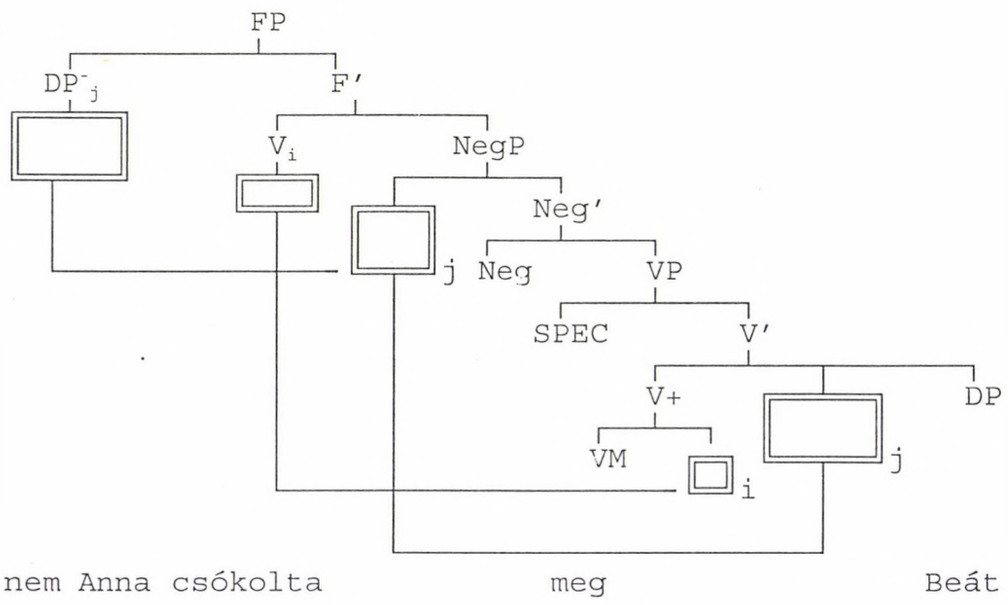


3.3. Phrase-negation

Phrase-negation can be deduced with the help of the negator-projection as follows: The DP, which should be negated, will be moved into SPEC-Neg, where it gets the negative-marker. With this act we have achieved that the phonological modul will interpret the DP correctly. However - as we have seen - in case of phrase-negation the focus-marker is also necessary, because there is a real banishment, so the verb has to be moved first into F to be able to distribute the focus-marker to the DP, which gets into SPEC-F.

(1) Nem Anna csókolta meg Beát.
It was not Ann who has kissed Beatrice.

(34)

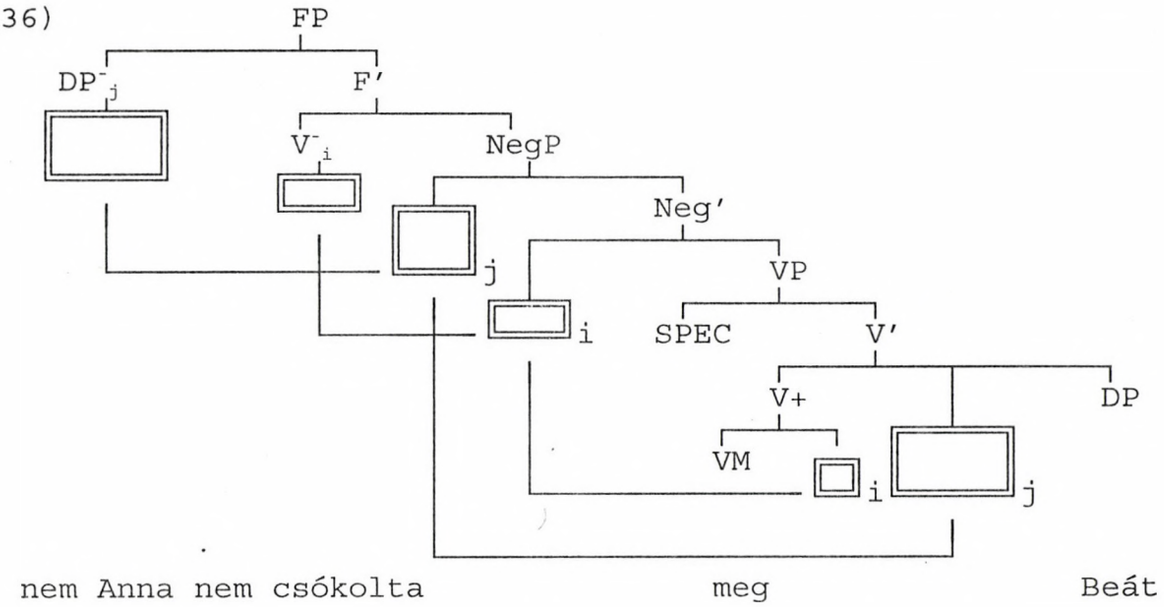


Another argument for this conception is that this model enables us to solve the problem of sentences which contain a phrase- and a sentence-negation as well, which we could not solve by the other two conceptions:

(35) Nem Anna nem csókolta meg Beát.
not Ann-NOM not kiss-PAST PREV Beatrice-ACC
It was not Ann who has not kissed Beatrice.

In case of (35) the route of the DP is the same as in (34), but the verb gets the negative-marker too, so it moves not directly into F, but first into Neg and just then into F:

(36)

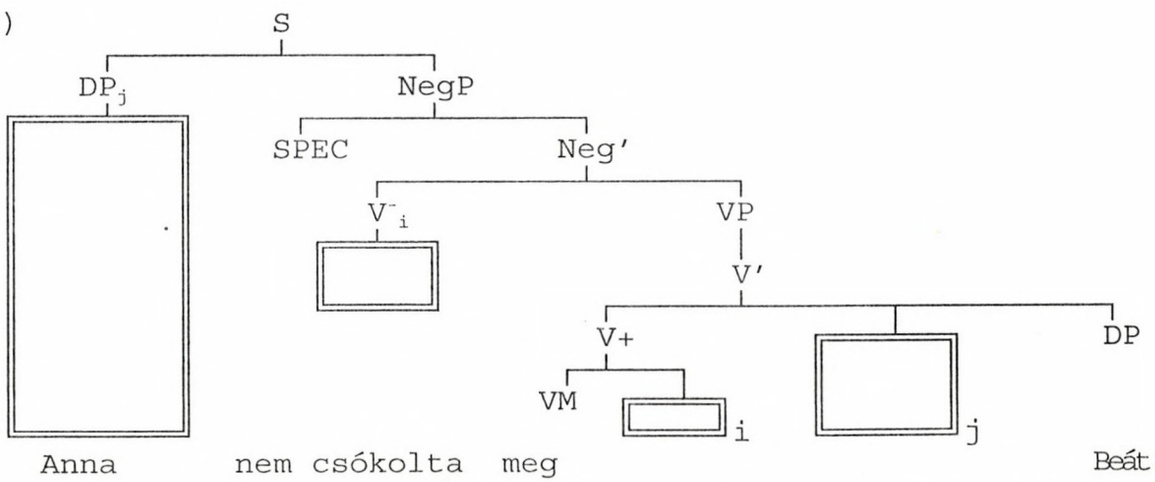


3.4. Sentence-negation

With this model we have also solved the dilemma we had in case of sentence-negation, namely, we do not need focus-projection in a netural sentence any more.

(4) Anna nem csókolta meg Beát.
Ann has not kissed Beatrice.

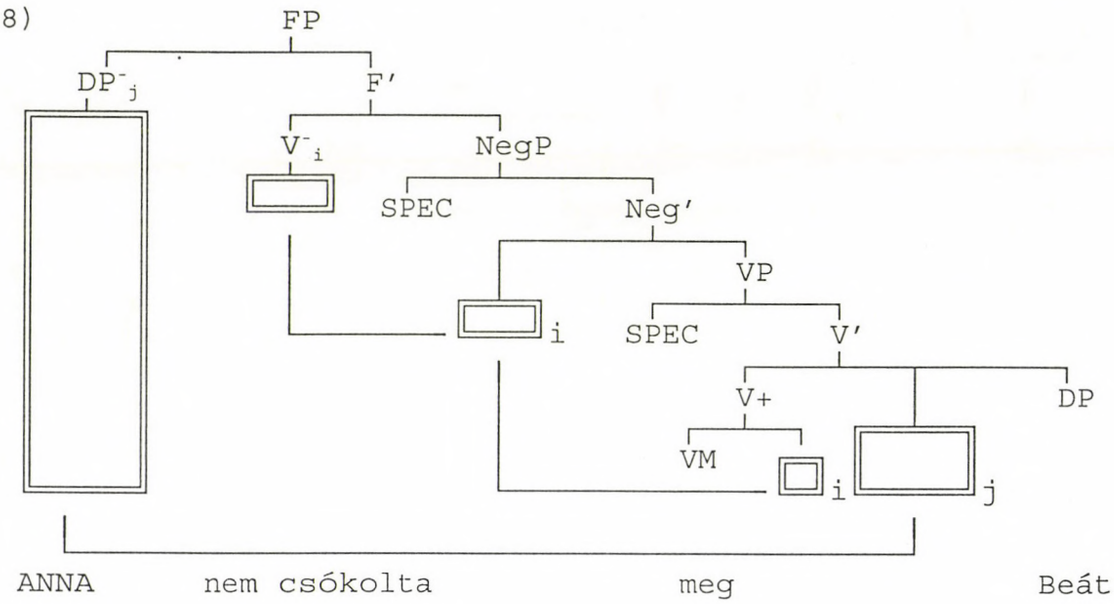
(37)



Furthermore, it is possible to represent the surface structures of sentences which contain a focused element and sentence-negation together, because there is no difficulty for the verb to move into Neg to get the negative-marker and then go on into F to distribute the focus-marker to a DP, which is directly moved into SPEC-F. So the surface structure of (24) is not a problem any more:

(24) ANNA nem csókolta meg Beát.
It was Ann, who has not kissed Beatrice.

(38)



3.5. Scope

Considering the question of scope by accepting negator-projection, we have to refuse the elegant idea which says: the operators precede and c-command their scopes already in the surface-structure in positions which precede the verb. It is all the more necessary, because the negator does not appear in the surface sturcture at all. Hence it follows that it cannot c-command any component. Therefore we have to define the interpretation of scope in another way.

It seems to be the simplest solution to interpret the component which has got the negative-marker as the scope of the negator. However - as we will see - this conception can be held only in some trivial cases:

(39) nem [Anna] csókolta meg Beát
not [Ann] has kissed Beatrice

That this is correct we have admitted already in (31).

In case of sentence-negation, the situation is different:

(40) Anna nem [csókolta] meg Beát
Ann-NOM not [kiss-PAST] PREV Beatrice-ACC

The sentence (4) ("Ann has not kissed Beatrice") declares that the pair <Ann, Beatrice> is not an element of the extension of the predicate KISSED. (40) is rather to be interpreted like: the pair <Ann, Beatrice> is an element of the extension of a predicate, but it is just not the predicate KISSED. Therefore it is true that the verb gets the negative-marker, but the scope of the negator is the argument list of the verb:

(40') megcsókolta(nem [Anna Beát])
kissed (not [Ann-NOM Beatrice-ACC])

To decide this question, a deeper semantic analysis is necessary, which cannot be part of this brief paper.

However, I would like to call attention to the fact that the question of the scope has to be examined in a wider sense of the word.

(41) Nem Anna csókolta meg Beát, hanem Cili.
not Ann-NOM kiss-PAST PREV Beatrice-ACC but Cecily-NOM
It was not Ann who has kissed Beatrice, but Cecily.

(42) Nem Anna csókolta meg Beát, hanem Bea Cilit.
not Ann-NOM kiss-PAST PREV B.-ACC but B.-NOM C.-ACC
It was not the case that Ann has kissed Beatrice, but Beatrice has kissed Cecily.

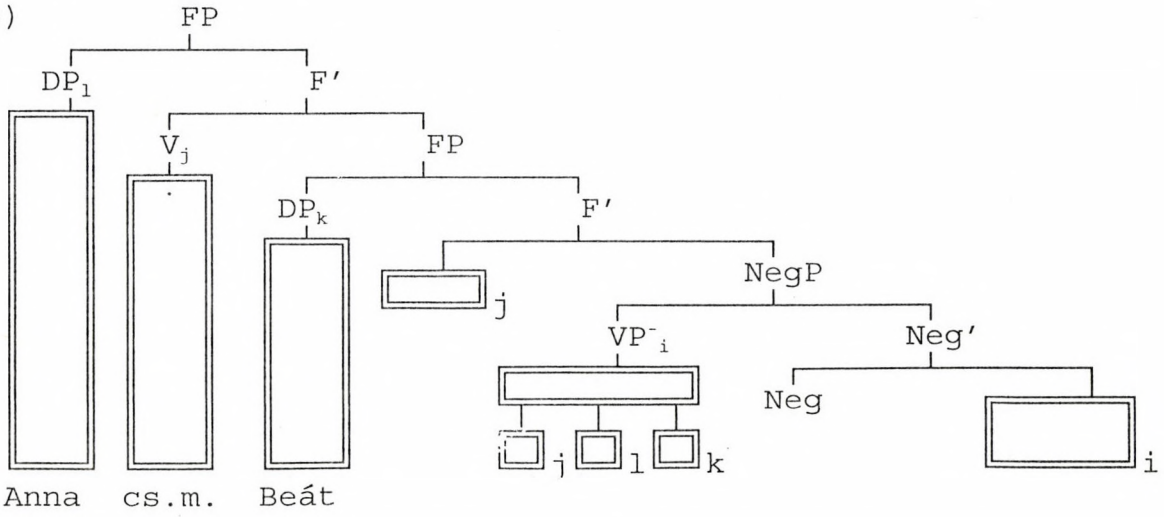
The structure of (41) is well known.

The structure of (42) is ambiguos: It is clear that in case of (42) we banish the pair represented by the two arguments from the extension of the predicate represented by the verb. It is a fact that the event described in the VP is not existent. It is a fact that another pair is an element of the extension of the predicate.

Solution 1:

So we negate the VP and we banish the pair, which is represented by the arguments of the verb. It follows from this that in (42) both arguments have to be moved into a SPEC-F and the VP must get the negative-marker.
Examine the structure of (42):

(43)



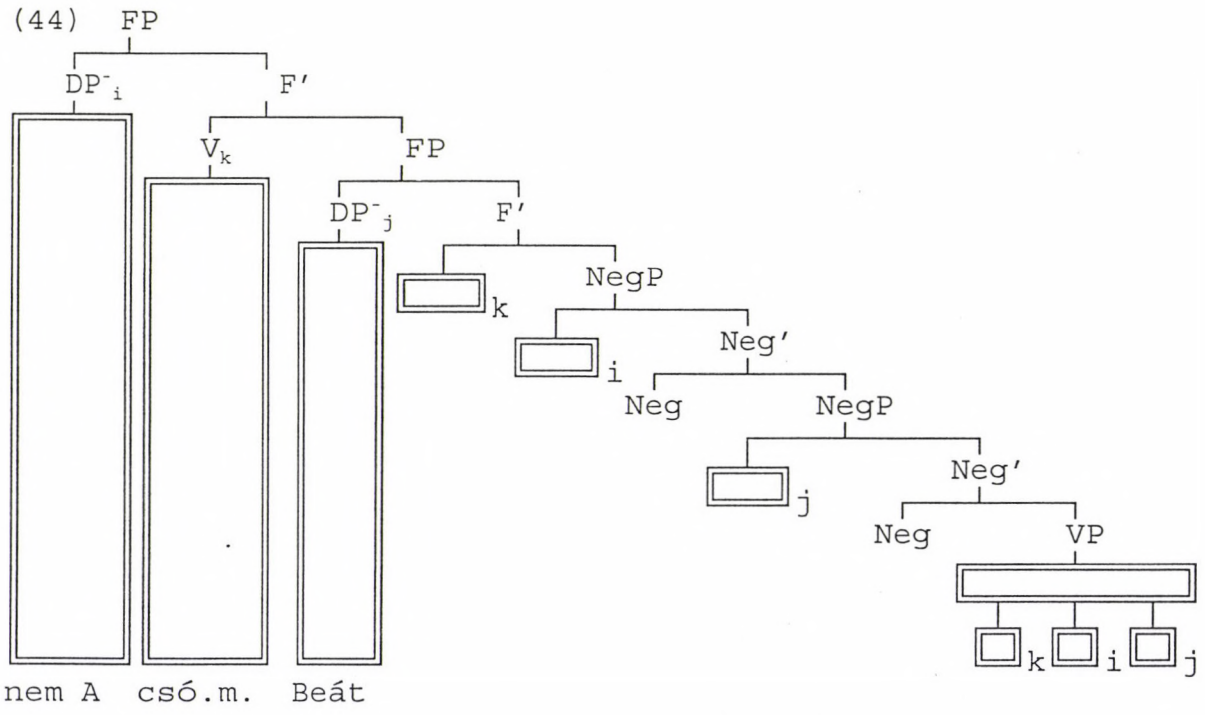
The question still remains: where will the phonological modul position the word 'nem', which is required to appear because of the negative-marker the VP has got. There are two premmisses:

1, the negator cannot be positioned behind the verb.
2, the verb itself has not got the negative-marker, so the negator cannot get directly before the verb.

So it can get only in a position which precedes the verb, but not directly.

Solution 2:

If we negate the VP in case the event it describes is not present, then we would have to negate the VP also in case of sentence-negation. As we have seen, the scope of negation is the argument list of the verb if the event described by the VP is not present. So the really consistent solution would be to negate both of the arguments. So we need a recoursive NegP. Both arguments would be moved into a SPEC-Neg. The verb must be moved into F - because of the banishment. And both of the DPs must move into SPEC-F. And because of the first premisse: "the negator cannot be positioned behind the verb" the second negator will be deleted:



References

1. Bródy Mihály: Some Remarks on the Focus Field in Hungarian In: Approaches to Hungarian 3. 1991.
2. É. Kiss Katalin: Az egyszerû mondat szerkezete. [The Structure of the Simple Sentence] ms.
In: Kiefer, Ferenc (Szerk.): Strukturális magyar nyelvtan 1. Mondattan. Akadémia Kiadó, Budapest 1992. pp. 79-179.
3. É. Kiss Katalin: Többszörös fókusz a magyar mondatszerkezetben. 1995.
4. É. Kiss Katalin - Szabolcsi Anna: Grammatikaelméleti bevezetô. [Grammar-theorethical Introduction] In: Kiefer, Ferenc (Szerk.): Strukturális magyar nyelvtan1.
Mondattan. Akadémia Kiadó, Budapest 1992. pp. 21-79.
5. Ruzsa Imre: Klasszikus, modális és intenzionális logika.[Classical, Modal and Intentional Logic]
Akadémia Kiadó, Budapest. 1984.
6. Ruszsa Imre: Logikai szintaxis és szemantika. [Logical Syntax and Semantics]
Akadémia Kiadó, Budapest. 1988.
7. Varga László: Observations on Negation in Hungarian and English. In: Acta Linguistica Academiae Scientiarum Hungaricae, Tomus 30 (1-2), pp. 67-96 1980.

Latin syllable structure

András Cser

0. Introduction

The purpose of this paper is to present a comprehensive model for the structure of syllables in Classical Latin. The value of such a model is, as I hope I shall be able to show, that it reduces a variety of observable facts to a very simple general pattern with the help of which all phenomena pertaining to the phonotactic organization of Latin words can be coherently explained and which also has remarkable relevance for some aspects of Latin morphology.

The paper is organized in the following manner: in section 1 the phoneme inventory of Latin is presented; in section 2 a taxonomy of consonant clusters is given; in section 3 the syllable template is introduced in a provisional form, then extended with extrasyllabic /s/; in sections 4 and 5 two further classes of phenomena are discussed, syllable contact and consonant manner assimilations at morphological boundaries, both of which will be seen to be closely linked to and governed by the sonority hierarchy, the basic organizing principle of many syllable-related regularities.

1. The phoneme-inventory of Latin

Classical Latin had the following phonemes[1]:

1) Vowels

i	u	i:	u:	ĩ:	ũ:
e	o	e:	o:	ê:	
a		a:		ã:	

Consonants

p	t	k	k^w
b	d	g	
f	s		h
m	n		
	l		
	r		
	j	w	

As can be seen, the Latin vowel system contained five short vowels, five long nonnasal vowels and four long nasal vowels.[2] In fact, nasal vowels are always derivable from V + N sequences, hence they are not to be posited as underlying segments in a generative analysis; since, however, they contrast with nonnasal vowels on the surface, they have been included in the above chart. An example of the three-way contrast:

2) /puʻella/ 'girl', /puʻella:/ 'girl ABL', /puʻellã:/ 'girl ACC'[3]

The vocalic sequences traditionally referred to as diphthongs can be shown (with surprising ease) not to be diphthongs but biphonemic sequences of a vowel in nucleus and a glide in coda[4]. Examples are:

3) *aes* /ajs/ 'bronze', *poena* /pojna/ 'punishment', *augere* /awge:re/ 'to grow'

All consonants, with the exception of /w/ and /h/, can occur as geminates, though many of them only at morphological boundaries:

4) *lippus* 'bleary-eyed', *siccus* 'dry', *aggerere* 'to amass', *pellis* 'skin', *peius* [pejjus] 'worse'.

2. Consonant clusters

The permitted combinations of consonants in Latin pattern as follows:

I) Initial

A) two-member

i) stop or /f/ + liquid: *tremunt* 'they tremble'
glaber 'bald'
frui 'to use'

ii) /s/ + voiceless stop: *spuere* 'to spit'
stare 'to stand'
squalor 'dirt'

iii) /s/ +/w/: *suadere* 'to persuade'
suauis /swa:wis/ 'sweet'

B) three-member

/s/ + vl stop + liq: *scribo* 'I write'
stridor 'hissing noise'
splendere 'to shine'

II) Intervocalic

A) two-member (geminates disregarded)

i) voiceless obstruents: *hospes* 'master'
uixit '(he) lived'
dictus 'said'

ii) /g/ + /m/: *agmen* 'march'
segmentum 'segment'

iii) sonorant + obstruent: *mentiri* 'to lie'

celsus 'high'

turba 'crowd'

caedo /kajdo:/ 'I cut'

iv) stop + liquid: *patres* 'fathers'

duplex 'double'

ebrius 'drunk'

v) liquid + nasal: *ulmus* 'elm tree'

cernere 'to see'

vi) liquid + /w/: *uoluit* /wolwit/ 'he rolls'

curuare /kurwa:re/ 'to bend'

vii) /m/ (or [ŋ]) + /n/: *somnus* 'dream'

(*dignus* [diŋnus] 'worthy')

viii) glide + sonorant: *aula* /awla/ 'court'

aeui /ajwi:/ 'age GEN'

B) three-member (never include geminates)

i) son + vl stop + /t/: *emptus* 'bought'

auctus /awktus/ 'grown'

ii) son + stop + liquid: *cultri* 'knives'

aegra /ajgra/ 'ill'

iii) son + /s/ + /t/: *faustus* /fawstus/ 'favourable'

instar 'as large as'

iv) vl stop + /s/ + /t/: *sextus* 'sixth'

depstum 'kneaded'

(C) four-member: *extra* 'outside'

plaustrum 'cart')

III) Final

A) two-member

i) son + vl obstruent: *pars* 'part'
laus 'glory'
sunt 'they are'
hunc 'this ACC'

ii) vl stop + /s/: *mendax* 'liar'
daps 'banquet'

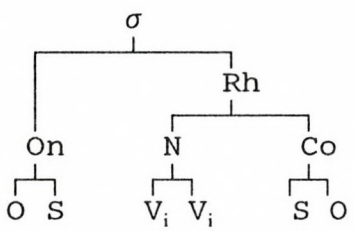
B) three-member (only monosyllables!)

son + vl stop + /#s/: *falx* 'scythe'
urbs [urps] 'city'
faex /fajks/ 'dregs'

3. The general structure of syllables

The apparent variety of consonant clusters can be promptly reduced to very simple regularities if we assume that a syllable in Latin can have the following maximal structure[5]:

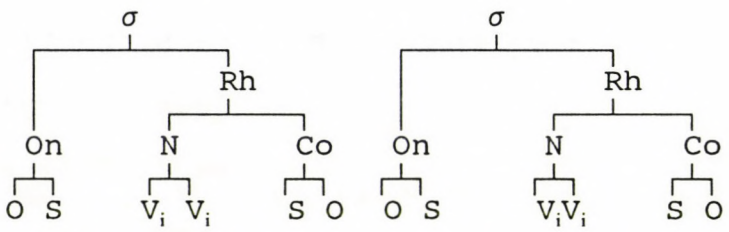
4)



A maximal syllable can thus have two timing slots in the Nucleus, which have to be attached to the same root node, two positions in the Onset, one for an obstruent and one for a sonorant, and the reverse in the Coda. Consonant clusters fit into the pattern in the

following way[6]:

5)



(I A i)	t	r	e			.		m	u			n	t
(I A iii)	s	w	a	a		.		w	i				s
(II A i)	h		o			s.	p		e				s
(II A ii)			a			g.		m	e			n[7]	
(II A iii)	k		a		j	.	d		o	o[8]			
(II A iv)	p		a			.	t	r	e	e			s
(II A v-vii)			u		l	.		m	u				s
(II A viii)			a		w	.		l	a				
(II B i)			a		w	k.	t		u				s
(II B ii)	k		u		l	.	t	r	i	i			
(II B iii)	f		a		w	s.	t		u				s
(III A i)	p		a		r	s							

Unforunately, there appear to be certain clusters that do not fit into the template given above. These are:

i) #{*s*}stop: *{s}tare, {s}tridor* (I A ii, I B)

ii) stop{*s*}stop: *dep{s}tum* (II B iv)

iii) stop{*s*}#: *dap{s}, urb{s}* (III A ii, III B).

They all include an /s/ between a stop and a word boundary or between two stops. At first glance, these /s/-es could only be

syllabified if we permitted two adjacent obstruents to be tautosyllabic. This, however, will soon turn out to be an incorrect solution.

First let us consider /s/+stop sequences. Unequivocal evidence shows that word-internal /s/+stop sequences are not syllabified as complex onsets: *hospes* 'master', *hostis* 'enemy', *Fuscus* PROPER NAME all have a heavy first syllable despite the short vowel in it. What this shows is that /s/ + stop is not a possible onset in Latin. Consequently, word-initial /s/, if followed by a stop, cannot belong to the onset.

Another argument comes from verbal morphology. Many Latin verbs form their perfective forms with reduplication[9]:

6) *tondeo* → *to-tondi* 'cut'
   cano → *ce-cini* 'sing'
   do → *de-di* 'give'

Reduplication is conspicuously absent in the case of verbs that begin with a complex (O+S) onset:

7) *tremo* → *tremui* 'tremble'
   suadeo → *suasi* 'persuade'
   floreo → *florui* 'blossom' etc.

However, some verbs that begin with /s/+stop do show reduplication:

8) *sto* → *s-te-ti* 'stand'
   spondeo → *s-po-pondi* 'promise'

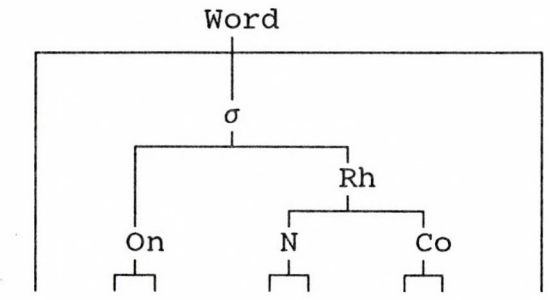
Also notice that the process of reduplication simply disregards the /s/ and copies only the stop.

These facts are enough to show that word-initial /s/, if followed by a stop, cannot belong to the onset. In theory, it could belong either to the word but not to any of its syllables, or to the first syllable as a kind of an appendix outside the onset, but there seems to be a strong argument against the latter view, which will also show the close relationship between the three problematic positions of /s/.

An apparently strange phonotactic rule of Latin is one that bans the occurrence of two hitherto unsyllabifiable /s/-es in a word (ie. word-initial before stop, word-final after stop, word-internal between two stops); that is, there are no words of the form */{s}pak{s}/, */{s}pak{s}ta/, */map{s}tik{s}/. This means that the distribution of these /s/-es is regulated not at the level of the syllable, but at the level of the word, ie. (and as follows from the previous arguments), these /s/-es are **extrasyllabic**[10]. As is to be expected, extrasyllabic /s/ freely combines with Onset and Coda /s/-es: *{s}pes* 'hope', *{s}pissus* 'dense' etc. Regular exceptions, in which two extrasyllabic /s/-es occur, are formations with the productive (feminine agent noun) suffix *-trix*: *speculatrix* 'female spy'[11].

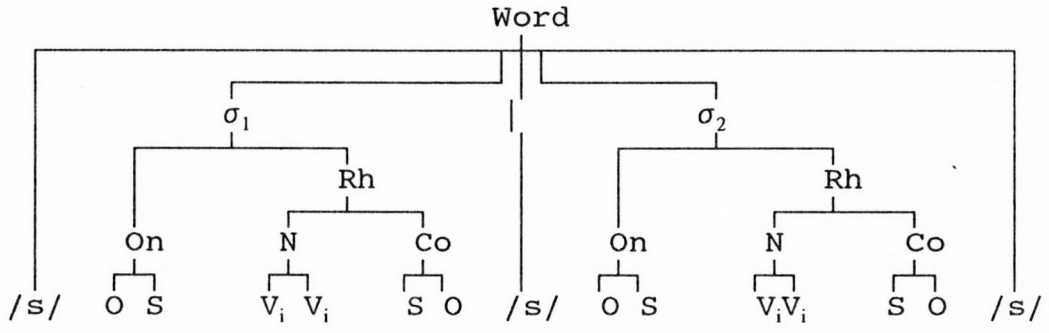
The following chart shows that, if we assume extrasyllabic /s/-es with the distribution described above, all the remaining clusters can be syllabified:

9) monosyllables



	/s/	O S	V_i V_i	S O	/s/
(I A i)	{s}	p	e e	s	-
(III A ii)	-	d	a	p	{s}
(III B)	-		u	r p	{s}

polysyllables



	/s/	O S	V_i V_i	S O	/s/	O S	V_iV_i	S O	/s/
(I A i)	{s}	t	a a		-	r	e		-
(I B)	{s}	t r	i i		-	d	o	r	-
(II B iv)	-	d	e	p	{s}	t	ũ ũ		-
(III A ii)	-	m	e	n	-	d	a a	k	{s}

A few points still remain problematic: the unexplained lack of branching codas before word-internal extrasyllabic /s/: **e(mp)*$_{Co}${*s*}*tus*, but cf. *e(mp)*$_{Co}$*tus*, *de(p)*$_{Co}${*s*}*tum* and *u(rb)*$_{Co}${*s*}; branching onset after word-internal extrasyllabic /s/ is only found in *extra* /ek{s}(tr)$_{On}$a:/, but cf. *{s}(tr)*$_{On}$*idor* etc. (Branching coda and branching onset can be adjacent, cf. *pla(us)*$_{Co}$*(tr)*$_{On}$*um*, so there

can be no simple syntagmatic restriction that bans four adjacent consonants, though such clusters are, in fact, infrequent.)

4. Syllable contact

In this section the exinsting syllable contact types are listed according to the major class of the consonants at the boundary (Obstruents, Nasals, Liquids and Glides):

10)

	O	N	L	G
O	hospes	(agmen)	(patres)	---
N	mentiri	(somnus)	---	---
L	celsus	ulmus	---	uoluit
G	caedo	poena	aula	aeui /-jw-/

Notes:

The *agmen*- and the *somnus*-type are much rarer than the other types; the *patres*-type actually does not normally divide its internal cluster between the two syllables, both consonants tend to be assigned to the second syllable. Since, however, such clusters are occasionally heterosyllabic, this type is included in the chart too. The only relatively frequent heterosyllabic clusters in which the sonority of the Coda is lower than that of the Onset is /lw, rw/.

5. Manner assimilations at prefix-stem boundary

Manner and place assimilations often take place at the boundary of a prefix and a stem. The assimilations are always regressive and often optional. Since here we are concerned with sonority

relations, place assimilations will be disregarded. In this chart it seemed useful to subdivide the class of Obstruents into Stops and Fricatives.

11)

	St	Fr	N	L	G
St	---	Yes (1)	Yes (2)	Yes (3)	No (4)
Fr	No (5)	---	[	(6)	]
N	No (7)	No (8)	---	Yes (9)	No (10)
L	No (11)	No (12)	No (13)	---	(Yes) (14)
G	No (15)	No (16)	No (17)	No (18)	---

examples and comments to (11):

(1) *ad+ferre → afferre* 'to take there'

(2) *ad+mouere → ammouere* 'to move there'

(3) *ad+leuare → alleuare* 'to lift up'

(4) *adiuuare* 'to help'

(5) *distinere* 'to keep apart'

(6) /s/ generally banned before voiced segments, /h/ in any coda, /f/ allowed in coda only in derived geminates (see (1) *afferre*)

(7) *continere* 'to keep together'

(8) *insidere* 'to sit on/in', nasals deleted before fricatives

(9) *con+rigere → corrigere* 'to correct'

(10) *iniuria* 'injustice'

(11) *perdere* 'to lose'

(12) *perspicere* 'to notice'

(13) *permanere* 'to stay long'

(14) untypical, the only example is *per+iurare* → *peierare* /pejje-/ 'to perjure oneself', but *periurare* as well

(15) *praeponere* 'to put before'

(16) *praeferre* 'to prefer'

(17) *praemittere* 'to send forward'

(18) *praelegere* 'to read out'

As can be seen, sonority considerations are conspicuously at work here: assimilation may take place if the sonority of the last consonant of the prefix is lower than the sonority of the first consonant of the stem; since assimilations are always regressive, this means that they eliminate those syllable contacts, where sonority relations are dispreferred.

NOTES

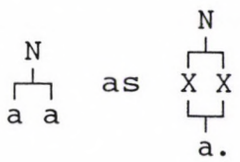
1.On the phonetics of Latin see Allen (1978), Leumann, Hofmann & Szantyr (1963-65), Niedermann (1953).

2. Normative spelling in the Antiquity did not distinguish long and short vowels.

3.The spelling of the forms is *puella, puella, puellam,* respectively.

4.I have done the analysis in detail in my thesis paper; in this paper only the major points will be hinted at.

5.I disregard certain questions throughout the paper, such as the restrictions that hold between nucleus and coda.
For ease of exposition I omit the X-tier in the charts; the representations must be interpreted accordingly, eg.



6.Evidence for the syllabification of intervoicalic consonants comes from versification in most cases, but also from the operation of various syllable-related processes, such as stress assignment. On these see Allen (1978), Steriade (1982) and (1988). On Latin syllable structure see also Panfilov (1973), though Panfilov's taxonomy has a number of serious flaws.

7.*-gm-* is, in fact, the only case in which a voiced obstruent is allowed to occur in Coda.

8.The syllable-based argument for the non-existence of diphthongs in Latin is that the sequences in question (*ae, oe, au*) are never followed by tautosyllabic sonorants, which shows that the glide in these sequences is itself the element that occupies the only Coda position reserved for sonorants.

Other arguments come from vocalic alternations, in which "diphthongs" do not take part and the non-appearance of /aw/ (as opposed to /aj/) word-finally, which is fully consistent with the general ban on word-final noncoronal consonants.

9.Vowel change may also occur, but this is now immaterial. The forms given below are PRESIMPFSING1 and PRESPERFSING1.

10.On extrasyllabicity see Goldsmith (1990); see also Törkenczy (1994) on Hungarian syllable structure, where the notion of appendix is also made use of.

11.The only real exception is *strix* 'owl'.

BIBLIOGRAPHY

Allen, W. S. (1978) *Vox Latina*. 2nd ed. Cambridge: Cambridge University Press.

Goldsmith, J. A. (1990) *Autosegmental and Metrical Phonology*. Oxford: Basil Blackwell

Leumann, M., J. B. Hofmann & A. Szantyr (1963-65) *Lateinische Grammatik*. Handbuch der Altertumswissenschaft 2.2.1-2. München: Beck.

Niedermann, M. (1953) *Phonétique historique du latin*. 3. éd. Paris.

Panfilov, E. D. (1973) *Fonologičeskije slogi klassičeskoj latyni. Čast' I: Odnosloz'nije slovoformi*. Leningrad: Izdatelstvo Leningradskogo Universiteta.

Steriade, D. (1982): *Greek Prosodies and the Nature of Syllabification*. Ph.D. Dissertation, MIT.

Steriade, D. (1988): 'Gemination and the Proto-Romance Syllable Shift' in: D. Birdsons & J. P. Montreuil (eds.): *Advances in Romance Linguistics*. Dordrecht: Foris.

Törkenczy M. (1994) A szótag [The syllable]. In: Kiefer F. (ed.) *Strukturális magyar nyelvtan II. Fonológia [A Structural Grammar of Hungarian II. Phonology]*. Budapest: Akadémiai Kiadó. 273-392.

Aniko Csirmaz

Swahili Possessive Structure and PP*

1 Introduction

Since Chomsky 1970, it has been widely accepted that nominal and clausal constructions share similar structure. The similarity is further emphasized by the assumption that there exists a predication relation within the genitive nominal structure itself:

(1) $[_{DP}$ D $[_{XP}$ [NP] X [YP]]],

where YP is predicated of NP, resulting in the English surface structures

(2)a YP's NP or

b NP of YP

as a crude approximation. In the approaches that involve the incorporation of a preposition, the category YP = PP will give rise to NP_{poss} in (2), as in den Dikken 1995b.

The category of the preposition *of,* apparently an inserted preposition, seems rather dubious, characterized as D/P or C by Kayne 1994 and as a nominal copular, the nominal counterpart of *be* by den Dikken 1995a.

In this paper I will argue that there is an embedded PP functioning as a complement of X in (1), X assumed to be Agr, following Szabolcsi 1983. In Swahili nominal possessives there is an obligatory particle, the equivalent of the English preposition *of,* which will be argued to be the dative P itself. Invoking Swahili PPs as well, the dative preposition will be shown to appear in PPs alongside the elements that were traditionally considered to be genuine prepositions. The dative Ps provide the complement DP with Case; and the same dative preposition is found heading *by*-phrases in passive constructions, again, providing Case.

*I would like to acknowledge the help of Marcel den Dikken, Géza Füssi-Nagy, and Mark Newson in preparing this article.

Swahili is a Bantu language spoken by a rapidly growing number of people in Eastern and Central Africa. Most of the speakers use Swahili as a *lingua franca,* though there are also a number of native speakers. Swahili is basically a non-configurational language, with a configurational DP, similarly to Hungarian. It makes much use of prefixes that are defined with respect to abstract noun classes. By the different prefixes marked in italics Swahili can express logically connected ideas: e.g. *m*-tu 'man-sg', *wa*-tu 'man-PL', *ki*-tu 'object-sg', *vi*-tu 'object-PL', *u*-tu 'humanity' share the root *-tu*. Prefixes are not obligatorily overt on nouns.
Numbers and genuine adjectives agree with the noun head, as in

(3) *m*-wavuli *m*-moja *m*-baya
umbrella one bad
'a bad umbrella'

Genuine adjectives are scarce in Swahili, and mostly the possessive construction is used to express modification (cf. 3.1). Verb agreement and tense appear as prefixes, while different aspect, including reciprocity, itinerant activity and passive are realized as suffixes:

(4) a- me- m- fung- ua *m*-lango
he- PAST- class2,SG- close- OPPOSITE door
'he opened the door'

I present a more detailed description of Swahili possessives and PPs in section 3.

2 Theoretical Framework

2.1 Antisymmetry

Kayne 1994 derives X-bar theory without stipulations by establishing the *Linear Correspondence Axiom* (LCA), that linearly orders any pair of nodes which share an asymmetric c-command relation. It follows from LCA that only one constituent can be adjoined to another one, complying with the binary branching restriction as well. Furthermore, only leftward movement and adjunction is allowed, adjunction involving

elements that are both heads or non-heads. I also assume Swahili to be a head-first language with respect to DPs, in the sense of Kayne 1994.

2.2 The existence of AgrP within DP

Szabolcsi 1983 established an AgrP as the complement of D based on the following Hungarian facts:

(5)a [$_{DP}$ az [$_{AgrP}$ én Agr háza-m]]
the I house-my
'my house'

b [$_{DP}$ az [$_{AgrP}$ ô Agr háza-0]]
the he house-his
'his house'

The agreement between the pronoun possessor and the suffix of the possessee suggests the existence of Agr, which I will henceforth suppose to be universally present in nominal genitival constructions. Thus *John's car* can minimally be assigned the structure in (6).

(6) [$_{DP}$ D [$_{AgrP}$ John Agr car]],

where Agr will be spelled out as *'s*.

2.3 Possessors as Dative PPs

Both Kayne 1994 and den Dikken 1995b derive clausal possession from the nominal one, although making use of different mechanisms.

(7)a the brake of the car/ the car's brake

b the car has a brake

To form (7b), Kayne 1994 suggests (7a) to be assigned the structure in (8), where IP is the nominal possessive construction. HAVE results by incorporating D^0 to BE.

(8) BE [$_{DP}$ D [$_{IP}$ DP I XP]]
└──┘

The structure suggested by den Dikken 1995b takes the alternation between possession and location, similar to (7) into account:

(9)a John gave the car *to Mary*

b John gave *Mary* the car

The dative PP is a locative predicate since it can undergo locative preposing, parallel to (10b), where the locative PP headed by *down* is preposed:

(10)a to Mary was given the car t

b down the hill rolled the baby carriage t

Based on these facts, den Dikken claims that double object constructions involve the incorporation of *to*. Similarly, possessors originate as complements of the dative preposition and clausal possession is derived by incorporating the head of the preposed PP into BE, resulting in HAVE, as in (11).

(11) BE $[_{PP}$ P DP$]_i$ $[_{XP}$ DP X $t_i]$

According to den Dikken, then, in all possessive constructions there is a dative PP involved having the possessor DP as complement, while no such projection is assumed by Kayne. Regarding the lowest phrase, labelled IP in (8) and XP in (11), it may tentatively be labelled AgrP, since we assumed it to exist universally.

Relations within AgrP are exactly opposite under the two contrasted analyses: according to Kayne's approach, the possessor appears in [Spec, AgrP], while under den Dikken's analysis it is within the complement of the AgrP. The different basic positions may prove relevant in establishing predication relations (cf. section 3.2).

Another possibility would be that constructions with a dative PP account for inherent possession, and the structure involving plain Agr for the 'idiomatic' ones:

(12)a the baby's beauty

b (that) beauty of a baby

Were that assumption true, there would be no way to account for the fact that structures like (12b), usually labelled *N of a N* constructions show the same range of phenomena that inverted predicates do (Bennis, Corver and den Dikken 1996). Consider (13a), illustrating the ungrammaticality of extraction of the subject of the inverted predicate, and (13b), showing

the same effect for N of a N construction.

(13)a **which neighbor* do you think that in front of the gate parked t?

b **a baby* which I consider [a beauty of (a) t]

Kayne's analysis does not give the desired results, because no predicate inversion is involved under his approach. Therefore the dative PP analysis will be assumed to hold generally for all nominal structures.

3 Possessive Structures and PPs in Swahili

3.1 Possessive Structures

The Swahili possessive particle, similarly to English *of,* Dutch *van,* and German *von,* among others, obligatorily follows the possessee and also bears a class prefix that agrees with it:

(14)a *ki*-tabu *ch*-a/ **w*-a/ **y*-a/ **l*-a *m*-ke
book of woman
'the woman's book'

b sungura *y*-a/ **w*-a/ **ch*-a/ **vy*-a *m*-toto
rabbit of child
'the child's rabbit'

Similarly, many ideas usually expressed by a noun modified by an adjective in Indo-European languages are a possessive construction in Swahili:

(15) *m*-tu *w*-a/ **y*-a/ **kw*-a akili
man of brain
'the smart man'

Facing the data above, it seems an attractive solution to unite the dative PP-hypothesis and the idea of AgrP as a complement of D:

(16) $[_{DP}$ D $[_{AgrP}$ kitabu Agr $[_{PP}$ P_{dat} mke]]]

The existence of the dative is justified in (14) and (15), the DP following the possessive particle being predicated of the noun preceding it. The basic position of DPs being determined, the question concerning the status of the possessive particle remains. There are at least four possibilities given the present assumptions: either it is a) D itself, having AgrP as its complement; b) a category X dominating DP, both requiring raising of the noun appearing in initial position; c) Agr itself; or d) the dative preposition heading the complement of Agr.

(17)a $[_{DP}$ kitabu$_i$ *D* [t$_i$ Agr $[_{PP}$ P mke]]]

b $[_{XP}$ kitabu$_i$ *X* $[_{DP}$ t$_i$ D $[_{AgrP}$ t$_i$ Agr $[_{PP}$ P mke]]]]

c $[_{AgrP}$ kitabu *Agr* $[_{PP}$ P mke]]

d $[_{AgrP}$ kitabu Agr $[_{PP}$ *P* mke]]

The first possibility involves raising of the possessee to [Spec, DP], which could account for the shared prefix established under Spec-head agreement. This approach is plausible based on the fact that all possessive constructions have a definite reading (witness the translation of the examples). To achieve indefinite reading, the overt indefinite article, *moja* must be inserted:

(18)a gari la *m*tu
car of man
'the car of the man'

b gari moja la mtu
car one of man
'a car of the man'

c gari la *m*tu *m*moja
car of man one
'the car of a man'

It should be noted that the possessive particle is also present when the structure has an indefinite reading. Not arguing for the suggestion presented here in detail, I would like to propose that the definite article that is present in neutral constructions is phonologically null. Demonstrative pronouns, that are an overt representation of D, allow optional raising of the predicate over D, both word orders being neutral:

(19)a *ki*-tabu [$_D$ hi-*ki*] *ki*-kubwa
book this big

b *ki*-tabu *ki*-kubwa$_i$ [$_D$ hi-*ki*] t_i
book big this

assuming the predicate originates to the right of D, as assumed previously. *Kitabu* may also be raised from a lower position, but this fact is not of immediate concern here. Possessive particles do not allow similar raising:

(20)a *ki*-tabu *ch*-a *m*-ke
book of woman

b **ki*-tabu *m*-ke *ch*-a
book woman of

Apart from inversion facts, a lower projection functioning as the possessive particle would also be advantageous for it would language-internally motivate the existence of the independently stipulated Agr or P. Furthermore, it would be consistent with widespread phenomenon of the cardinal 'one' being phonologically identical to the indefinite article (such as Hungarian *egy,* Spanish *un,* German *ein*). Accepting the possessive particle to have the category D would force us to assume that the cardinal *moja* functions in Swahili as a -- presumably -- unique converter, changing the definite article to an indefinite one. Also, in
DPs not involving possession, a null D would have to be stipulated anyway:

(21) ni- na- la nyama ya nguruwe
I- PRES eat meat of pig
'I am eating (some specific piece of) pork'

In summary, analysing the possessive particle as D would not account for lack of predicate inversion and possible definite interpretations of non-possessive DPs and would not be consistent with a universally observable tendency.

Raising the possessee to the specifier would also be necessary for X, X dominating DP. However, this solution involves the stipulation of an additional projection for which there

is no motivation. Trying to keep necessary structure to a minimum, no unmotivated projections will be assumed.

Agr^0 was stipulated to ensure agreement between the Hungarian possessor and the possessee, the latter bearing a suffix agreeing with the former. The situation in Swahili is quite different; it is not the possessor, but the possessive particle that agrees with the possessee. Let us suppose that Agr is basically the same in the two languages, and there may be an additional, higher category above AgrP that determines the superficial differences. If the possessive particle is Agr, the shared prefix is accounted for by Spec-head agreement.

The possessive particle may also have the category P, being the dative preposition. We will return to this possibility in 3.2 to review it in more detail. It should be noted that incorporation of P to Agr is a necessity, to ensure agreement between the possessee *kitabu* in Spec, AgrP, and the possessive particle. Although obligatory raising may seem less economical than the previous possibility, the particle being Agr, other motivations for this path will be presented later on.

According to the discussion above, the possessive particle may either have the category Agr or P, all involving raising or incorporation. The following sections will shed some light on the competing categories.

3.2 Swahili PPs

Prepositional phrases involve an obligatory possessive particle intervening in Swahili, where the particle agrees with the preposition:

(22)a	juu	*y*-a/ **w*-a/ **ch*-a/ **mw*-a	*ki*-ti
	above/ on	of	chair
	'on the chair'		
b	*ku*-toka	*kw*-a/ **y*-a/ **w*-a/ **l*-a	somo
	about/ from	of	lecture
	'about the lecture'		

An obvious description of this structure would be to treat *juu ya* and *kutoka kwa* as complex prepositions, spelled as two words for some reason or another. This approach to the

phenomena would only restate the problem in a different way, for the question regarding agreement would still linger. As it has been shown in the previous section, the possessive particle is not capable of determining the prefix it bears alone. The preposition preceding the particle must do so; then it is more economical to assume the preposition to require a particle then establish each preposition and particle as a complex unit.

It is regarding prepositional constructions as sharing the structure with possessive constructions that we can exclude Kayne's proposal for the nominal genitival construction in (8). Assuming AgrP reflects the predication relations within it, it can be maintained that these are valid with DP-complements of Agr (with P adjoined). Claiming that the same is true in PPs, that is, *kutoka* is predicated of *somo,* its complement, is arbitrary, to say the least.

Another possible approach is to assume that the preposition itself (*juu* and *kutoka*) originates as head of the dative PP and then is obligatorily raised to Spec, AgrP to ensure agreement. This would prove the hypothesis that the possessive particle is in fact Agr.

Accepting the suggestion above runs counter several morphological and thematic facts. The semantic differences could not be appropriately accounted for, since all 'prepositions', that is, *juu*, *kutoka*, and so on, would originate as dative Ps and raise to Spec, AgrP. They would then be predicted to have the same meaning.

It is also observable that 'prepositions', or *Prep-markers,* as I will henceforth call them, determine the form of the possessive particle the same way nouns do, although the variation of possible prefixes is very limited. Infinitives receive the prefix *ku-*, and behave just like nouns in all respects; this is not surprising, since infinitives received a noun class prefix. I do not wish to claim it is the prefix that makes the infinitive a noun, it only serves as
a diagnostic. The same prefix may appear on the Prep-markers *kufika* and *kutoka*:

(23)a *ku*-fika *kw*-a *m*-ji
in of city
'to the city'

b *ku*-toka *kw*-a gari
about/ from of car
'from the car'

Other Prep-markers bear no overt class prefix and must be followed by *y-a,* the particle bearing the prefix of class 5, the class including most nouns without overt prefixation. (24a) lists such nouns, while (24b) includes a few such Prep-markers.

(24)a *gari* 'car'; *sungura* 'rabbit'; *chai* 'tea'; *namna* 'kind, type'

b *tangu* 'from'; *chini* 'under'; *mbali* 'far away'

These facts suggest that Prep-markers are in fact nouns and occupy the same position the possessees do:

(25) $[_{AgrP}$ kutoka Agr $[_{PP}$ P gari]]

In this section, it has been shown that the most likely candidates for the category of the possessive particle are Agr and P_{dat}, since the Kaynian analysis of nominal possessives has been excluded because of predication facts in PPs. The prepositional phrases that consist of a noun-like P- marker, a particle, and a noun have essentially the same structure as possessives involving two DPs. This assertion leads to the supposition that Prep-markers and possessor DPs behave similarly with respect to Case checking, too. That is, the particle is essentially required for checking Case, whatever its category is.

4 Passivization and the dative PP-analysis

4.1 Facts about passive

Two remarks are appropriate concerning section 3.2. First, there is a third type of particle occurring with Prep-markers, where the particle bears no class prefix. Such a particle accompanies the Prep-marker *karibu*, meaning 'beside', 'close to':

(26) karibu *n*-a/ **w*-a /**ch*-a/ **y*-a/ **l*-a *ki*-sima
beside of well
'beside the well'

Providing an answer to this problem is likely to prove easier after considering the paradigm in (27). The examples of passive construction with 'by'-phrases allow only a restricted set of prefixed particles; this phenomenon constitutes the second problem. Only the particles *na* and *kwa* are allowed, all other forms are ungrammatical. Hopefully, the answer to one of the puzzles will also account for the other.

(27)a Shaaban a- na- pend- wa
Shaaban he- PRES- love- PASSIVE
n-a/ *kw*-a/ **w*-a/ **y*-a mama
by mother
'Shaaban is loved by Mother'

b *ma*-tunda ma- ta- la- wa
fruit-PL class3,PL- FUT- eat- PASSIVE
n-a/ *kw*-a/ **y*-a wewe
by you
'the fruits will be eaten by you'

Apparently, in these constructions the sole function of the particle is to check Case of the agent. Since the sequence kwa/ na + DP is not embedded within a DP, the existence of Agr is not enforced under the present assumptions. Also, no agreement is possible because neither a Prep-marker nor a DP is allowed to precede the sequence:

(28)a **ma*-embe ma- ta- la- wa *ma*-pera
mango-PL class3,PL- FUT- eat- PASSIVE pear-PL
y-a/ *n*-a/ *kw*-a wewe
by you
'the mangoes and the pears will be eaten by you'

b **ma*-embe ma- ta- wek- wa
mango-PL class3,PL- FUT- put- PASSIVE
karibu *kw*-a/ *n*-a yeye
beside by him
'the mangoes were put by him beside himself'

The particle + DP sequence is then simply a PP that is also present embedded in AgrP in nominative possessives, P being the particle and the agent its complement:

(29) $[_{PP}\ [_{P}$ kwa/ na$]\ [_{DP}$ wewe$]\ ]$

Since the particle appearing as the equivalent of the English *by* in passives and *of* in possessives are identical, the category of the possessive particle is P_{dat}; and thus the question whether the possessive particle is of category Agr or P is tentatively resolved.

4.2 Further evidence for P_{dat}

In unmarked constructions, the object of a ditransitive verb may, but its goal must be expressed by a prefix only. If the goal receives emphasis it follows the object DP as complement of *kwa*:

(30)a ni- li- (ki-) ku- pa kitabu
I- PAST- (class4,SG-) you- give book
'I gave you the book'

b ni- li- ku- pa kitabu *kw*-a/ **n*-a/ **y*-a wewe
I- PAST- you- give book to you
'I gave the book *to you*'

The 'preposition' (in traditional terms) that appears in (26a) is a dative preposition. It has the same form the possessive particle (the term includes the particle found in *by*-phrases) that was assumed to originate in PP. This finding supports both the dative PP-hypothesis and the presumption formed in section 3.2.

The possessive particle *kwa*, similarly to most Prep-markers of Swahili, is not only equivalent to the preposition *to*. Other meanings include *meanwhile, according to, because of,* and *onto*. Some examples are given below:

(31)a m- na- zungumza *kwa* Kiswahili?
youPL- PRES- chat in Swahili
Are you talking in Swahili?

b ni- me- anguka *kwa* wewe
I- PAST- fall because of you
'I have fallen because of you'

The variation in Indo-European, or more specifically, English prepositions assigned to each Prep-marker and the possessive particle itself only reveals the fact that the partition of the spatial and temporal concepts is different. This is neither unique nor surprising; consider, for example, the fact that Hungarian has two different suffixes to express movement and state:

(32)a az asztal-*ra* rakom a virágot
the table- on put-I the flower-ACC
'I put the flower on the table'

b az asztal-*on* van a virág
the table-on is the flower
'the flower is on the table'

The relevant fact showing that *kwa* is the overt realization of the dative P is that no other form of the possessive particle and no Prep-marker can appear in that position:

(33)	ni-	li-	ku-	pa	kitabu	(*karibu/ *katika)	
	I-	PAST-	you-	give	book	beside	in
	kw-a/ **w*-a/ **n*-a/ **y*-a			wewe			
	to			you			
	'I gave the book to you'						

The analysis presented shows that the possessive particle has the category P, that obligatorily moves to Agr to satisfy agreement requirements. The exact nature of this movement is not clear; it will be supposed that it is adjunction to Agr. This operation is permitted by LCA and is also compatible with the discussion in a Minimalist vein of section 5. Agr only serves in Swahili as a landing site for P; the movement itself is driven by the necessity to establish agreement between P and the possessee DP.

A problem is still to be answered, namely, the question why *kwa* functioning as a dative P bears the class prefix of the class of infinitivals. The other equivalent of *by*, the particle *na* also poses a similar problem, namely, why it bears a prefix that is unattested in other contexts, be it on nouns, adjectives, verbs or demonstratives. In addition, while *kwa* can be identified to bear a class prefix, the same cannot be said of *na*, for no overtly prefixed nouns, adjectives or verbs have a prefix *n-*. This problem is to be discussed in the following section.

5 Underspecified forms and the Minimalist Program

5.1 Case Checking and Agreement

Deviating somewhat from the assumptions outlined in section 2, in this section I will assume the Minimalist mechanism of feature checking under a Spec-head configuration. Thus not only agreement in terms of noun class prefixes, but Case checking also requires that a head with one of the above features enter into a Spec-head relation with the checkee. Movement is feature movement, where the differences between the outcome and trigger of movement of formal and lexical features follow Categories and Transformations of Chomsky 1996. Within this framework, all movement is feature movement. Covert or phonologically unobservable

movement involves only formal features (Case, category and ϕ features), while in overt movement lexical features are pied-piped with the formal ones. No stray features are allowed, that is, if one formal feature moves, all must move. Feature checking takes place in a Spec-head configuration, where features of the head and that of Spec are checked against each other, and non interpretable features are eliminated, although they may still be observable for morphology.

In the following sections, I will assume the basic structure derived by the end of section 4. I will also retain the previous labelling of projections, although this is not strictly in accordance with the presently assumed checking mechanisms. For example, Case checking is supposed to be realized by an Agr node. Case of the subject DP is checked by AgrS; however, the Case feature of AgrS is provided by V. In terms of the analysis of possessives, this would mean that an Agr node has to be stipulated on top of P, where P would adjoin Agr to check Case of DP:

(34) $[_{AgrP}\ DP_j\ Agr\text{-}P_i\ [_{PP}\ t_i\ t_j]\ ]$

The AgrP in (34) is clearly distinct from the AgrP assumed earlier; the former one is a functional projection inserted to check Case. To remove confusion, the Agr of Szabolcsi 1983 would have to be renamed, say, to *Con,* from *concord.*

Since the issues sketched in the previous paragraph are not of crucial importance in the following argumentation, they will not be represented or referred to; however, they should be kept in mind.

5.2 Case and Raising in Specific Constructions

In the possessive *DP of DP*- type constructions, the possessive particle, the head of the dative PP raises to adjoin to Agr. Agreement in terms of class prefixes is accounted for, since it was supposed that the possessee DP sits in [Spec, AgrP]. The feature that defines the prefix choice, say *gender feature* (a ϕ feature) is checked in this configuration. This gives the desired result: Case of the possessee DP will be checked by a higher category, the one that was traditionally assumed to assign Case to DP; the exact mechanism will not be considered. The possessor, however, is in need to check its Case feature; hence, it raises to [Spec, PP]. In

section 4, it was suggested that P is capable of checking Case: it checks Case of the possessee DP, and since the feature Case is not interpretable, it is eliminated. The rest of the formal features raise to adjoin Agr.

Exactly the same procedure is applied in *PP*s, the constructions that contain a Prep-marker in place of the possessee. Again, P checks Case of its complement, as is required for DPs. The Prep-marker, however, has no Case-checker; traditionally, the need does not even arise, for prepositions themselves can check the Case feature of DPs. Under the analysis put forward here, the Prep-marker behaves just like a noun, but has no Case features to be checked; in other words, of the formal features, it has only ϕ features, among them, gender feature and categorial one -- hence, the variation between *kwa* and *ya* follows. Raising P to incorporate into Agr satisfies the need to check the gender features of the Prep-marker and no successive raising is needed. The Prep-marker also has lexical features; these ensure the different meanings associated with different Prep-markers.

Finally, the equivalents of *by-phrases* indicate that raising of the DP complement of P is, in this case, covert. Recall that the structure assigned to *kwa wewe* 'by you' was:

(35) $[_{PP}$ e $[_{P}$ kwa] wewe]

To check its Case, *wewe* must raise to [Spec, PP]. Since no higher projection was assumed, if movement were overt, the DP would precede the head.

5.3 Feature composition and feature movement

Since the dative PP is essentially the same in possessive and passive constructions, raising the possessor in *DP of DP* constructions should also be covert, accepting the conclusion at the end of the previous subsection.

Raising the possessor to Spec, PP must precede raising of P itself, for otherwise the Case feature could not be checked, no stray features being allowed. Thus, P must also raise covertly. Under the approach of Categories and Transformations of Chomsky 1996, this equals to saying that P bears lexical features; if it had none, the formal features of the DP would obligatorily pied-pipe the lexical ones of the DP itself, enforcing it to be spelled out in Spec, PP.

By analogy, Agr also has lexical features and thus only the formal features of P are obliged to raise to the checking position. A similar result was achieved in Zwart 1996, where he claims that Holmberg's Generalization is also valid for Dutch. In Dutch, scrambling is possible without overt movement of V. Zwart's conclusion therefore is that shift of V is required as shift of formal features. Since only the formal features of V raise to AgrO and higher functional positions such as T and AgrS, it follows then that Agr nodes, at least in these languages, have lexical features.

5.4 Underspecified Forms

The gender feature of P, being feature of a head, is non interpretable. This means that the derivation will crash if it is not checked. Checking this feature is problematic within by-phrases, since no projection Agr was assumed that was used in other instances to check gender, and the construction also lacks a DP to check it against, the DP being the possessee or the Prep-marker in the other cases.

There are basically two possibilities to overcome this problem. One of them would be to stipulate the existence of two different Ps; one with and one without a gender feature. The objective of this paper is to show that the two Ps are, in fact, identical; consequently, it would be desirable to reach a different conclusion. The other possibility is that there is in fact an Agr node on top of PP in by-phrases. P raises to Agr, as before, and there is a phonologically null element in Spec, AgrP. Agreement with this null element is realized on the Agr-P complex as *kwa* or *na,* the alternation unaccounted for. In other words, if there is no overt DP or Prep-marker sitting in Spec, AgrP, then Agr-P will have an underspecified phonological realization.

It should be noted that Agr needed not to be stipulated to exist in by-phrases in the previous sections, for the necessity to check features did not arise.

6 Conclusion

In this paper, it has been shown that the analysis proposed by Kayne has certain shortcomings in describing the structure of nominal possessives; among others, it does not account for phenomena shared by inverted predicates in N of a N constructions and apparent violation of predication relations in Swahili PPs. This led us to adopt Dikken's dative PP hypothesis, where possessors originate from the complement position of *to*. Furthermore, Agr was also stipulated as existing universally in DPs.

Swahili possessive constructions involve a possessive particle that follows and agrees with the possessee. The category of this particle seemed rather dubious at first sight. By examining PPs in Swahili, that were shown to consist of a Prep-marker + particle + DP sequence, the particle resulted in having certain preposition-like features, Prep-markers being defective nouns. Further evidence that the particle is in fact the dative preposition was gained from by-phrases in passive structures and PPs expressing GOAL. The latter constructions obligatorily involved some primitive, an agreementless form of the particle.

In the last section Case and gender feature checking were discussed in Minimalist terms, the latter defining the form of the noun class prefix. It was also suggested that the particles lacking agreement with an overt element agree with a null element, enforcing the existence of Agr even in by-phrases.

References

Bennis, Hans, Norbert Corver and Marcel den Dikken. 1996. *(A)symmetries in Predication and the Meaning of Structure.* Paper presented at GLOW 19, Athens. Ms., HIL/CLS

Chomsky, Noam. 1970. 'Remarks on Nominalization'. In Chomsky, ed. 1972. *Studies on Semantics in Generative Grammar.* The Hague: Mouton

Chomsky, Noam. 1996. *The Minimalist Program.* Cambridge, MA: MIT Press

den Dikken, Marcel. 1995a. *Copulas.* Paper presented at GLOW 18, Tromsø. Ms., Vrije Universiteit Amsterdam/ HIL

den Dikken, Marcel. 1995b. *Particles. On the Syntax of Verb-Particle, Triadic and Causative Constructions.* Oxford: Oxford University Press

Kayne, Richard. 1994. *The Antisymmetry of Syntax.* Cambridge, MA: MIT Press

Szabolcsi Anna. 1983. The possessor that ran away from home. *The Linguistic Review* 3:89-102

Zwart, Jan-Wouter. 1996. *Morphosyntax of Verb Movement. A Minimalist Approach to the Syntax of Dutch.* To be published by Kluwer, Dordrecht

English Linguistics Department

ELTE

Ajtósi Dürer Sor 19., H-1146 Budapest, Hungary

aniko@math-inst.hu

THREE QUESTIONS ON THE STATUS OF IMPERSONAL SENTENCE STRUCTURES

Dudás Kálmán

As a definition, a sentence structure is impersonal if

a it contains a verb in the 3sg form and
b there is no nominative subject in the sentence /other than a dummy/

Semantically, the structure expresses

a the non-dynamic /stative/ action and the participant's non-consciousness

Ex 1 Taedet nos sermonis illius.
bores us speech of that person
"That person's speech bores us."

b an indefinite and/or generic agent /a conscious participant

Ex 2 Es wird getanzt.
It is danced. "There is dance going on."

These two semantic points read as a OR b.

This short paper examines some European nominative languages /Latin, German, English, and Hungarian/ and answers 3 questions tersely owing to lack of space.

A. Is the above impersonal structure /from now on: impersonality/ uniform with regard to its syntanctic behaviour of a possibly mixed set of verbs? Or, where have all the subjects gone?
We might assume that there are superficially /surface-/ identical impersonalities that happen to have different - though interlinked? - morphological and/or lexical properties that lead to their common lack: lacking an agreeing subject.

B. Should impersonality be regarded as an isolated construction in the overwhelming number of personal constructions?

C. What can one call an /inherently/ impersonal verb?
/A terminological question. The morphological answer is straightforward but we wonder which verbs are never used otherwise /= in personal construction/?

To begin with, impersonality and subjectlessness seem to appear at the same time /see Appendix/.

3 Pudet me stultitiae meae.
is ashamed I-Acc. stupidity my-Gen/Gen
"My stupidity brings shame on me / I am ashamed of my st."

But there is an obvious subjectful case:

4 Decet fortiter loqui.
suitable bravely speak-Inf. "It's suitable to SUBJECT speak bravely."

A better view is that non-nominative subjects are impossible to agree with any verb and the verb takes a /default/ 3SG. /There can be other default forms in some other language./
A nominative subject is missing if there is

1 no subject /syntactically/ at all.

2 a subject but that is not eligible for agreement.

Consider the appendix examples:

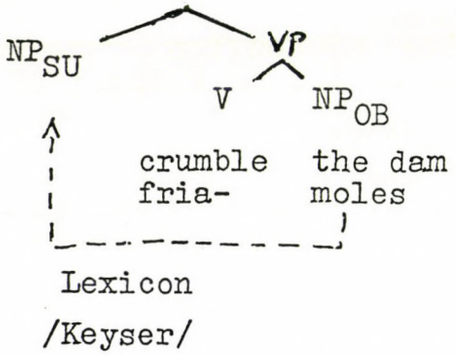
5 __Virrad. ___Pudet me malevolentiae tuae.

6 Mir graut /es/ vor ihm. Es graut mir vor jenem Mann.

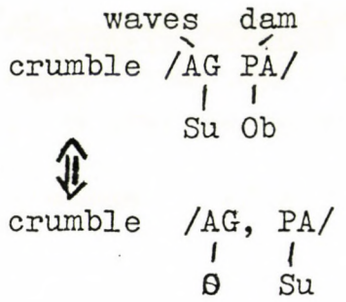
The German surface subject is a dummy. Comparing 5 and 6 we can find a +/- pro-drop distinction /Hung. and Latin are +proddrop./ Pro-drop languages do have an agreeing pro subject that is recoverable from the verb morphology or comtext. This is not the case here, for instance, for the Latin example you can find a 3Sg subject only which is in conflict with the 1Sg accusative semantically, moreover, pudet will become a 3-place predicate, or we have a superfluous argument thematically, etc. Expletive dummy it or es is forced by the extended projection principle in classical GB, here we treat them as formal subjects and lexical nominative subjects /that may be pronominal, of course/ count as subjects for our definition /p.1/.

The argument number of verbs is quite varied /Appendix 2/. This diversity is characteristic of unaccusative verbs /=the subject of an intransitive verb shows the absolutive properties of the object of a transitive verb in that language system/. There is always a "transitive pair" in the lexicon /App.1/.

This unaccusative group is a simple illuminating example of thematic-syntactic relations and also a group in close relationship with passives /discussed later/. The syntactic conversion in which the dam/moles emerges as the subject can be modelled by NP movement /Keyser on the lexicon,1993/ or by lexical redundancy rules /Bresnan,1982//Komlósy,MSM/.①,②



/Keyser/



/MSM/

/⟨⇒ means that the domain of this relation is narrower than any of the two verb classes which might be defined by purely formal properties of entries on either side of the relation/

Latin does have the passive voice /-tur/ as the morphological device, unlike English, and the reminiscent passive derivation by NP movement /eg. Jaeggli,1986/. But this conversion is just as much restricted lexically as in English or Hungarian. In this latter morzsol and /össze/morzsolódik /V and Ve/ have the same word root and the morphological change takes place in the stem /which is supposed to be longer than the root/ and this process is by no means productive, cf.

7 gyógy-ít : gyógy-ul "heal" olva-szt : olva-d "melt"
pörg-et : pör/ö/g "twirl" szegény-ít : szegény-ed
"make poor" "become p."

This last example has a free morpheme /szegény "poor"/ but that does not mean that -ít and -ed fits in with free morphemes. /Further morph. details in MSM/ The existence of a "re-link" rule is of great importance: it changes the quality of the complement relations linked with the argument slots /i.e. the syntactic functions of the complements expressing the arguments/. Lexical redundancy rules operate to form reflexives in Hung.: fésül "comb" - fésül-köd "comb himself".

A most affected experiencer subject is present in the PU-DET type thematic structure /German graut, schaudert/, with

stative verbs. The Hungarian examples /below/ show that arguments bearing a stimulus thematic role affect the whole argument structure and nominalization is impossible.

8 Sajnálom az apámat. Cf.*az én sajnálásom
I regret my father-Acc. I- regretting-
/pro drop/ gen. nom.poss.
"my regretting"

Péter ismeri a térképet
P.-Nom. knows the map-Acc. ??? az apám sajnálása
my father regretting-
gen. nom.poss.

*Péter ismerése ??? A térkép ismerése
P-Gen. knowing-nom.poss. map-Gen. knowing-nomp.

What is impossible in the Hungarian nominalization system is impossible in some Latin and German finite verb system cases. /In fact, the number of these verbs is 5 in Latin and about the same in German./ An example with similar meaning to that of Latin/German:

9 Péter tart Évától
P-nom. fear Éva-ablative
experiencer stimulus

*Péter tartása *Éva tartása
*Évától tartás/a/

We would like to focus on the fact that these are stative verbs and they lack agent or patient thematic role.

Along the lines of Legendre et al. /1993/ optimality model about case and voice systems, the prominence of either the patient or the agent /as they termed it, "aP or Ap"/ will affect the whole sentence structure, either of them may surface as subject in case$_1$ /=in nominative languages: nominative case/. The structure that can emerge as output from the low-prominence /"aP"/ agent and high-prominence patient input in nom/accusative languages has passive structure with an agent demoted to an oblique NP /by-agent in English, ab-agent in Latin,etc./ or a non-surfacing one. But the case of "ap" /both low-prominence/ is not clearly dealt with. It seems /above exx./ that for a PUDET verb's experiencer there is a way of remaining demoted even if the stimulus argument is similarly demoted: in fact, it is the ergative/absolutive pattern /eg.

Eskimo/ in which "aP" gives a structure where the agent is an oblique-case or surface-absent NP and the patient argument receives an absolutive case like intransitive arguments. /I have simplified their description, the morphological case is the NP's possibility of realizing abstract case, other means being word order or verbal cross-reference./ Mutatis mutandis, the experiencer /= a quasi-patient/ and the stimulus /= a quasi-agent/ behave in this way in Latin/German isolated ergative structures with finite verbs and in Hungarian with nominalizations. We have examined the PUDET type where it is the thematic structure that determines subjectlessness.

Quite differently,the ITUR type follows the nom/acc predictions in Legendre's model as far as arriving at the classical passive subject demotion and passive morphology is concerned. Two-argument verbs /paret/ may undergo this conversion as the one-argument verbs more frequently do /see App.1/. As there is no patient argument /as object/ which, consequently, cannot receive the case from the verb, following Burzio's generalization, the subject does not have to receive a theta role /although it may receive one/. Therefore, theta criterion and case filter are satisfied.

This group always has the active-voice agentive counterpart in which there is a nominative subject /=a personal construction is available/:

10 Fugitur.
Run-3SG pass.
"There is running."
"Some people are running."

Marcus fugit.
Marcus-nom.sg. Run-3SG act.
"Marcus is running."

11 Es wird getanzt.
It dance-3SG pass.
"There is dancing."
"Some people are dancing."

Peter tanzt.
P-nom.sg. dance-3SGact.
"P. is dancing"

It follows that ITUR can only be a syntactic derivational group. Opposed to -dynamic PUDET, one may dub this argument structure as "supple". Morphologically this suppleness is reflected in the passive forms /wird+past partic.//esse+p.p./, the marker of the indefinite or/and general subject.

Unaccusatives do not undergo this impersonal passivization:

Lacus inarescit. * Inarescitur.
/see Appendix 1/

And neither do personal passives:

12 Nullus homo a muliere excipitur.
13 *A nullo homine a muliere excipitur.
14 Es wird niemand von der Frau geküsst.
15 *Es wird von der Frau von niemand geküsst. /see App.1/

These verbs describe -conscious, -dynamic actions, and we do not have to assume that "suppleness" is some special feature, but it is exactly the +conscious and +dynamic combination, which is in the semantic characterization of the verb anyhow, and determines some semantic properties of the sentence, for instance, its aspect.

In English impersonal passives could be well-formed on general principles but "ap" /=low-prominence agent and patient/ is not allowed for some reason, andt imtransitive verbs cannot be passivized:

16 *It was danced.

In Hungarian there is no productive passivization at all, and like NP movements are doubtful, object raising is limited to hagy "let" /a possible control construction/:

17 János hagyja Pétert aludni.
János-nom. lets Péter-acc. sleep-infinitive

Also, subject raising is limited to the látszik/tűnik type /if not raising, a possible two-place verb/:

18 János aludni látszik.
János-nom. sleep-inf seems

A possible treatment for Hungarian lexical pairs as olvaszt/olvad /see earlier/ is lexical rules between the items. We will address this property of Hungarian at the end of the article.

An interesting subcase is Es tanzt sich /App. 1/. It is a quasi-passive where after agent demotion the analytic verb+reflexive pronoun construction replaces passive morphology, in other respect it is the ITUR type. On the one hand, several verbs do not follow this pattern:

19 *Es singt sich.
It sings itself.

It is hard to find cross-linguistic equivalents with the same meaning:

2o *Se salţat. *Táncolja magát. *It dances itself.
Itself dances. Dances itself.

On the other hand, some transitive verbs may have this form:

21 Solche Sachen sagen sich.
Such things-nom. say-plur itself
plur.
"Such things are said."

In German this is a general way to form unaccusatives:

22 schlängeln "wind sg"; schlängeln sich "wind, snake"
schließen "shut sg" ; schließen sich "close /around/"

So this seems to be a lexicalized conversion whether into impersonal structure or unaccusative /=best treated in the lexicon/.

In fairness to linguistic variety, there is one more subtype with some examples only, Es klopft /App.1/, an always derived impersonal construction:

23 Wasser stiebt.
Water sprays /out/. "Water sprays /out/."

24 Es stiebt.
It sprays. "It drizzles./There is drizzling of rain."

These verbs have the supple property and active morphology to express an indefinite subject. They may be close to weather verb /see below/ as 24 shows. Further examples:

25 Es klopft an die Tür.
It knocks at the door. "There is knocking at..."

26 Es klingelt.
It rings. "There is ringing./Someone is ringing."

At the moment we know less about these types than needed. The gradience towards impersonal weather verbs suggests that KLOPFT is a less conscious type semantically than ITUR. Like in the case of PUDET you experience the action that happens to you. Again, being some verbs that can follow KLOPFT pattern we should treat them lexically as the TANZT SICH type.

There remained one more reason for not having a surface subject. Provided the semantic argument structure holds no argument /complement slot/ at all in the lexical representation, some morphologically regular verbs /with finite forms in all tenses and moods/ cannot allow any subject other than a dummy:

27 Ningit. Es schneit. Havazik. Havazna. Havazott.
/see App.1/

In Hungarian no participial, gerundial, or deverbal noun forms can be derived as evidence for the absent subject. For possible objections to subjectlessness see MSM Ch.5 and Komlósy/1994/.

To sum it up, a subjectless syntactic structure may arise

from a thematic structure similar to those of erg/abs languages /PUDET/. There are few isolated examples in Latin and German. Another cause of the emergence of a subjectless surface structure is a syntactic operation by which the conscious agent of the verb is demoted and this loss in the argument number is accompanied by morphological change /into passive/: **ITUR.** Thirdly, TANZT SICH and KLOPFT have active morphology otherwise belong to this group. For a change, the last NINGIT type has an argumentless lexical structure that strictly demands /"entails"/ that the syntactic structure be subjectless. In fact, PUDET with its (-agent,-patient) thematic structure, and ITUR with the supple structure have a similar determining semantic structure. It has been shown that it is the agent and patient thematic roles that play a distinguished part in impersonality, even with their absence /PUDET,NINGIT/. Their significance has turned out to be crucial in such different fields of syntax, morphology, or semantics as compound formation /Kiefer,93/ or aspects /Kiefer: MSM, Bende Farkas,95, Kálmán,95, Maleczki,95/.

One group has been left for discussion. If there is a syntactic subject—unlike the types we have seen—and still it takes no nominative /see the paragraph below example 4, point 2/. This can only happen if the subject is not ruled out by the Case filter: if it is not a NP. Some verbs with basically modal meaning /necessity, possibility, suggestion, etc./ have a propositional argument:

28 Illik az öregeknek köszönni.
29 Oportet sero ad me venire.
3o Es ist mir gelungen, meinen Freund in Hamburg zu treffen /see App.1/

The matrix verb has one complement expressed by a finite clause or infinitive clause. In Hungarian this infinitival clause contains a dative DP and an infinitive /or an empty subject if the NP has moved out/, just like the possessive structure and mechanisms proposed by Szabolcsi /1983/,MSM. This view has been subject to debate, for instance, in the course of PhD lectures at the Research Institute for Linguistics, Budapest, from February to June 1996. Here we would like to point to the following fact of the whole verb system. If

this dative DP is an experiencer argument, the OPORTET type has two arguments out of which the propositional one can surface as subject. This impossibility of having an experiencer subject is not supported by the fact that even the PUDET type is absent from Hungarian, all experiencers being realized as surface subjects without exception. For an opposite view /two-argument treatment/ see Komlósy /1994/. The exact movement of this element, the dative DP, and a possible pro in the structure is out of the scope of this article although some details are shown in the Note /point 3/. In parenthesis, a related case problem, the origin of the accusative of the DP in question in Latin has not yet been resolved:

31 Dicitur Caesarem venire.
"It is said that C. is coming." /see App.1/

We put this aside for a subsequent paper on non-overt categories and cases.

Some comprehensive remarks are in order here. Firstly, almost always intransitive verbs occur in impersonality. This follows from the absence of a patient argument, and the exceptional PUDET, which is transitive, has an experiencer in the accusative. Secondly, agent demotion is characteristic of ITUR with a related morphological passive. Thirdly, stative and non-conscious action verbs do not have, by definition, supple structure and cannot take a passive form /PUDET,OPORTET,NINGIT/.

To conclude, number alternation of syntactic arguments occur WITH or WITHOUT accompanying morphological change. Morphological change happens WITH or WITHOUT a change in the connection between syntactic functions /subject is in focus for our examination/ and thematic roles. And lastly, this syntactic-thematic link may change WITH or WITHOUT re-link /see earlier/. For this latter WITHOUT option the ITUR type, with a simple agent demotion, is a good example. /Another is Hungarian causative, <u>húz</u> "pull"-<u>húzat</u> "make sb pull", which has not been discussed./

Imagine a grammatical system without this flexibility of the above WITH/WITHOUT options, instead, it has the sole WITH option. We arrive at a system in which every above property

includes the other /in the set theoretical sense/. Then what is "available" for this system as far as the impersonal structures are concerned is the two most universal types, the semantically argumentless NINGIT, and OPORTET that avoids NP filter anyhow. Other types, ITUR, PUDET, TANZT SICH, KLOPFT must be missing. To express passive, reflexive, or medial structures is impossible syntactically.* But these semantic relations will be borne by an intricately developed double lexical item system connected by redundancy rules /in the LFG model/ AND/OR word order complexity. This imaginary system is almost identical with that of Hungarian.

Out of the three questions A and B have been answered at the proper points. Briefly,

A/ There are four major argument configurations that automatically lead to a non-nominative subject and thus agreement is impossible;

B/ Impersonality deeply roots in the whole verb system, in fact, no impersonal structural feature whatsoever is to be stipulated for them in a model.

We have identified the OPORTET type that can only be used impersonally for morphosyntactic reasons, the NINGIT type that can be only be used impersonally for thematic reasons, and in Latin and German the PUDET type, impersonal only, for thematic reasons Ágain/ among the verbs that are used in impersonal structures. These three types are impersonal verbs /the answer to question C/.

I am grateful to András Komlósy for his comments on an early sketch of this article.

* by syntactic derivational rules

APPENDIX - 1

A UNACCUSATIVE SENTENCES

A1 Lacus inarescit.
Lake-nom.sg. dries out-3sg.act.
"The lake is drying out."

Inarescitur.
Dries out-3sg.pass
"There is drying out."

A2 Waves crumble the dam.
Hullámok morzsolják a gátat.
wave-nompl. crumble-Pl.3.active
dam- acc.sg. def.

Undae molem friant.
wave dam crumble
nom.pl. acc.sg. 3pl.active

The dam crumbles.
A gát összemorzsolódik
dam crumble+ódik
nom.sg. 3sg.active,-def.

Moles friatur.
dam crumble
nom.sg. 3sg.pass.

A3 A kenőcs gyógyítja a sebet.
Ointment heals wound
nom.sg. 3sg.def. acc.sg.
"The ointment heals the wound."

A seb gyógyul.
wound gets healed
nom.sg. 3sg.indef.
"The wound is healing up."

B IMPERSONAL SENTENCES

B1 Es wird getanzt.
It aux. dance
3sg. past partic.
nom.sg.
"There is dancing./
People are dancing."

Pugnatum est.
fight- aux
nom. sg.neutr. 3sg.
past partic.
"There is fighting./
People are fighting."

Sic itur ad astra.
So go-3sg.pass. to star-acc.pl.
"This way one can get to the stars."

Paretur legibus.
obey-3sg.pass. law-dat.pl. "One obeys laws."

Nullus homo a muliere excipitur.
No man-nom.sg by woman-abl.sg. receive-as-guest:3sg.pass.

*A nullo homine a muliere excipitur.
By no man-abl.sg.

Es wird niemand von der Frau geküsst.
It aux nobody by woman kiss-past p.
3sg.

*Es wird von der Frau von niemand geküsst.
It 3sg by woman by no one kissed-past.p.

Transl.: "No man is received by the woman."
"There is receiving by the woman /by the man/."

"Nobody is kissed by the woman."
"There is kissing by the woman /by nobody/."

B2 Es tanzt sich.
It dance-3sg.act. itself-acc.sg. "There is dancing."

B3 Es klopft (an die Tür).
It knock-3sg.act. "There is knocking (at the door)."

B4 Pudet me malevolentiae tuae.
cause shame I-acc.sg. malevolence- your
3sg.act. gens.sg. gen.sg.
"I am ashamed of your malevolence."

Es graut mir vor jenem Mann.
It abhor I- with some man.
3sg. dat.sg. "I abhor some man/
Mir graut /es/ vor ihm. him."
with him

Mir / Mich schaudert vor ihm.
I-dat.sg I-acc.sg shudder-3sg with him.
"I am shuddering because of him."

B5 Havazik. Ningit.
Snow-3sg.act. Snow-3sg.act.

Es schneit. Havazna.
It snow-3sg.act. "It snows." snow-condit.3sg.act.

Virrad. Havazott.
Dawn-3sg.act. "It dawns." snow-past.3sg.act.

B6 Illik az öregeknek köszönni.
Suit-3sg.act. old man-dat.pl. greet-inf.
"It becomes one to greet old people.

Oportet sero ad me venire.
Must-3sg.act. late to I-acc come-inf.
"One must come late to my home."

Es ist mir gelungen, meinen Freund in Frankfurt
It aux I-dat. succeed my friend-acc.sg in F.

zu treffen
to meet-inf. "I succeeded in meeting my friend in F."

Dicitur Caesarem venire.
Say-3sg.pass. Caesar-nom.sg. come-inf.
"It is said that Caesar is coming."

APPENDIX - 2

Verbs and argument structures /possible variations/

1	V /x,y/ -- V /y/	loosen,crumble; kiköt,terem "put into port", "grow,bear"
2	V /x,y/ -- Ve /y/	fugare "make sb run" megfutamít fugere "run" fut
3	V /x,y/ -- V /x/	x read y -- x read
4	V /x,y/ -- Ve /x/	
5	V /x,y/ -- Vp/y /	x say y -- y is said
6	V /x/ -- Vp /Ø/	tanzen -- wird getanzt; klopft
7	V /x,y/ -- V /x,x/	x praebet y "x shows y" x praebet se "x proves to be..."
		x kämmen y "x combs y" x kämmen sich "x combs himself"
8	V /x,y/ -- V /y,y/	x sagen y "x say y" y sagt sich "people say y"

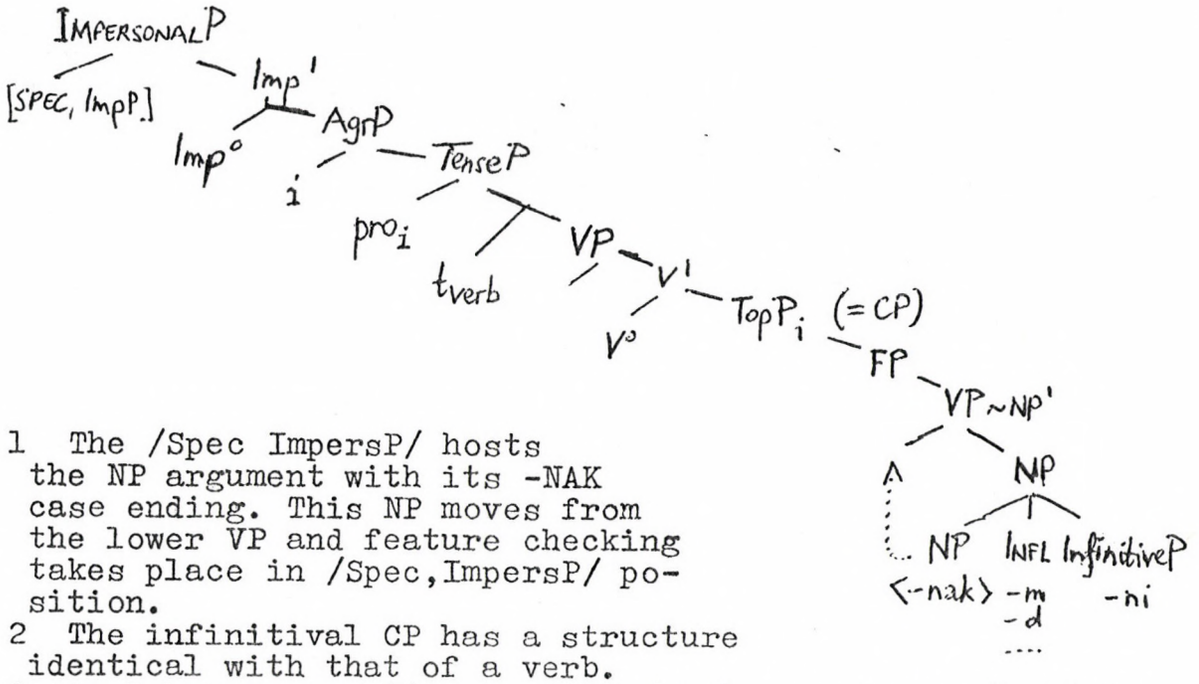
Ve = morphological change in the stem
Vp = passivized form
x,y = arguments /syntactically/
Ø = no syntactic argument
-- = connected in the semantic argument structure /no conversion is indicated/

NOTES

1 This article tends to refine a lexicalist point of view of impersonality. Nevertheless, the words "convert","demote", or"link" are not used terminologically in favour of, for instance, an LFG theory. Unaccusative=ergative.

2 This is a slightly formal model. The stress has been laid on the three questions on the basis of a corpus of c. 2oo sentences in 4 languages /a sample of which is shown in App.1/. A coherent formal model is yet to be processed.

3 A possible structure is a reshaped model of Benedicto /1995/. Here is a glimpse owing to lack of space:



1 The /Spec ImpersP/ hosts the NP argument with its -NAK case ending. This NP moves from the lower VP and feature checking takes place in /Spec,ImpersP/ position.

2 The infinitival CP has a structure identical with that of a verb.

3 AgrP has no specific meaning /number or person/: it is only a minimal expression of a pronominal head; its content is an i index /=the index of the CP argument of the verb/.

4 The expletive pro is coindexed with the CP argument.

5 The NP argument can move on from /Spec,ImpersP/ to the FP or TP projection /see Bródy,1995/.

Bibliography

Baker,M. /1988/, Incorporation: A theory of grammatical function changing. Chicago: Univ. of Chicago Press.

Bresnan,J./ed.//1982/, The Mental Representation of Grammatical Relations. MIT Press, Cambridge /Mass./; London.

Bródy,M. /1993/, Theta theory and arguments, LI 24, 1-24.

Burzio,L. /1986/, Italian Syntax. Dordrecht: Reidel.

Dowty,D. /1991/, Thematic Proto-Roles and Argument Selection. Language,Vol.67, No.3.

Benedicto,E. /1995/, Mne ne čitaetsja: Relativized Modality, Datives and Reflexive Suffixes. In: Penn Working Papers in Linguistics,Vol.2,No.2.

Carrier-Duncan,J. /1985/, Linking of thematic roles in derivational word formation. LI 16, 1-34.

Clark E. and Clark H. /1979/, When Nouns Surface as Verbs. Language Vol.55, 767-811.

Grimshaw,J. /1990/, Argument Structure. Cambridge /Mass./: MIT Press.

Hale,K. and Keyser,S.J. /1991/, On the syntax of argument structure. Lexicon Project Working Papers,32, MIT.

Jackendoff,R. /1987/, The Status of Thematic Relations in Linguistic Theory. LI 18, No.3

Jaeggli,O. /1986/, Passive, LI 17: 587-622

Kiefer,F./1993/,Compounds and Argument Structure in Hungarian.

In: I.Kenesei /ed/:Approaches to Hung.Vol.4.JATE,Szeged.
Keyser,S.J. and Hale K. /1993/, On Argument Structure and Lexical Expression of Syntactic Relations. In: The View from Building 20, Cambridge /Mass/, MIT Press.
Keyser,S.J. and Roeper,T. /1984/, On the middle and ergative constructions in English, LI 15, 381-416.
É.Kiss,K. /1986/, A személyragos főnévi igeneves szerkezetről /On the infinitival construction with personal suffixes/. Magyar Nyelv LXXXII, 393-4o8.
É.Kiss,K. /1989/, Egy főnévi igeneves szerkezetről /On an infinitival construction/. Általános Nyelvészeti Tanulmányok XVII, 153-169.
É.Kiss,K. /1987/, Configurationality in Hungarian, Dordrecht: Reidel, and Budapest: Akadémiai.

Komlósy,A. /1994/, Complements and Adjuncts. In: Syntax and Semantics, Vol.27 /The Syntactic Structure of Hungarian/, 91-178, Academic Press Inc.
Legendre,G.,Raymond W. and Smolensky P. /1993/, An Optimality-Theoretic Typology of Case and Grammatical Voice Systems. In: Proceedings of the 19th Meeting of the Berkeley Linguisitics Society
Marantz,A.P. /1984/, On the Nature of Grammatical Relations. Cambridge /Mass./: MIT Press.
Perlmutter,D. /1978/, Impersonal Passives and the Unaccusative Hypothesis. In: Proceedings of the Berkeley Ling. Society 4, Univ. of California
Resetsky,D. /1992/, Zero Syntax. Vol.1: Experiencers and Cascades. MIT Press.
Roca,I.M./ed.//1992/, Thematic Structure: Its Role in Grammar. Berlin: Foris.
Shopen,T./ed.//1985/, Language Typology and Syntactic Description.Vol.1: Clause Structure, Cambridge Univ.Press.
Szabolcsi,A. /1983/, The Possessor that Ran Away from Home. The Linguistic Review 3, 89-1o2.
Szabolcsi,A. /1991/, Suppressed or PRO subjects? The Argument Structure of Event Nominals in Hungarian. In: Approaches to Hungarian.Vol.3, JATE,Szeged.
Williams,E./1980/, Predication, LI 11, 2o3-38.
Williams,E./1981/, Argument Structure and Morphology. The Linguistic Review 1, 1, 81-114.
Zubizarreta M.-L. /1987/, Levels of Representation in the Lexicon and in the Syntax. Dordrecht: Foris.
In Approaches to Hungarian.Vol.5., JATE Szeged /I.Kenesei/:
Bródy,M. /1995/, Focus and Checking Theory. ApprH.5.29-45.
Bende-Farkas,Á. /1995/, Prefixation and Discourse. ApprH.5. 193-221.
Kálmán,L. /1995/, Definiteness Effect Verbs in Hungarian. ApprH.5.221-243.
Maleczki,M. /1995/,On the Definiteness Effect in Hungarian. ApprH.5.261-285.
In Magyar Strukturális Nyelvtan/I:Mondattan /A Structural Grammar of Hungarian/I: Syntax/, Akadémiai,1992:
Komlósy A. Régensek és vonzatok /Predicators and Arguments/, 299-528.
Kiefer,F. Az aspektus és a mondat szerk. /The Aspect and the Structure of the Sentence/ 797-886.

On Complements of Complements in Hungarian

Szécsényi Tibor

1 Goals

In the present paper Hungarian verbs having an infinitival argument, such as *akar* 'want', *kell* 'must', *szeret* 'like' etc. are to be examined. These verbs can be divided into two groups based on the following criterion: if they need a verb carrier[1] they function as auxiliaries, if they do not, they serve as main verbs (Kálmán et al., 1989):

(1) a. *Péter futni akar.*
P to-run wants
'Peter wants to run.'
b. *Péter el fog menni.*
P away will to-go
'Peter will go away.'
c. *Péter szeret úszni.*
P likes to-swim
'Peter likes swimming.'

In sentences (1a) and (1b) *akar* 'want' and *fog* 'will' are auxiliaries *futni* 'to run' and *el* 'away' are verb carriers. The verb *szeret* 'like' in sentence (1c) functions as a main verb having no verb carrier.

In the present paper I argue that these groups of verbs treat not only their own complements as complements, but also the complements of their infinitival complements. This is a well-known fact about the subject of the infinitive but we have to prove it for the rest of the complements. Therefore I am going to use the term *the complements of the infinitive* meaning *all complements of the infinitive but the subject.* Also, I am going to use the term *arguments of the verb* meaning *the complements and adjuncts of the verb.*

To define whether an argument is a complement of a verb or not we use the omission test proposed by Komlósy (in Komlósy, 1992, p. 316):

Omission test[2]: a constituent serves as a complement in a given sentence structure in case it is a complement which cannot be omitted.

(2) a. *Péter adni akar Marinak egy könyvet.*
P to-give wants to M a book-acc
'Peter wants to give Mary a book.'
b. **Péter adni akar Marinak*
c. *Péter szeret Marival találkozni.*
P likes M-with to-meet.
'Peter likes meeting Mary.'

---

[1] A verb carrier is that constituent of the senteces which lies in front of the verb immediately in neutral sentences. In non-neutral sentences it is moved behind the verb.

[2] There is another omission test (Komlósy, 1992, p. 318): in case an *x* argument licences the appearance of another argument that can be omitted, *x* is to be regarded as the (optional) complement of the regent.

On Complements of Complements in Hungarian

Szécsényi Tibor

1 Goals

In the present paper Hungarian verbs having an infinitival argument, such as *akar* 'want', *kell* 'must', *szeret* 'like' etc. are to be examined. These verbs can be divided into two groups based on the following criterion: if they need a verb carrier[1] they function as auxiliaries, if they do not, they serve as main verbs (Kálmán et al., 1989):

(1) a. *Péter futni akar.*
P to-run wants
'Peter wants to run.'
b. *Péter el fog menni.*
P away will to-go
'Peter will go away.'
c. *Péter szeret úszni.*
P likes to-swim
'Peter likes swimming.'

In sentences (1a) and (1b) *akar* 'want' and *fog* 'will' are auxiliaries *futni* 'to run' and *el* 'away' are verb carriers. The verb *szeret* 'like' in sentence (1c) functions as a main verb having no verb carrier.

In the present paper I argue that these groups of verbs treat not only their own complements as complements, but also the complements of their infinitival complements. This is a well-known fact about the subject of the infinitive but we have to prove it for the rest of the complements. Therefore I am going to use the term *the complements of the infinitive* meaning *all complements of the infinitive but the subject.* Also, I am going to use the term *arguments of the verb* meaning *the complements and adjuncts of the verb.*

To define whether an argument is a complement of a verb or not we use the omission test proposed by Komlósy (in Komlósy, 1992, p. 316):

Omission test[2]: a constituent serves as a complement in a given sentence structure in case it is a complement which cannot be omitted.

(2) a. *Péter adni akar Marinak egy könyvet.*
P to-give wants to M a book-acc
'Peter wants to give Mary a book.'
b. **Péter adni akar Marinak*
c. *Péter szeret Marival találkozni.*
P likes M-with to-meet.
'Peter likes meeting Mary.'

---

[1] A verb carrier is that constituent of the senteces which lies in front of the verb immediately in neutral sentences. In non-neutral sentences it is moved behind the verb.

[2] There is another omission test (Komlósy, 1992, p. 318): in case an *x* argument licences the appearance of another argument that can be omitted, *x* is to be regarded as the (optional) complement of the regent.

d. **Péter szeret találkozni.*

What is left to be decided is whether the complements of the infinitive can be regarded as arguments of the matrix verb. If they can, they are complements as well, for they cannot be omitted.

2 Lexical background

In case of NPs an alternative way to decide whether they are the complement of a given regent or not is to check whether it gets a semantic (thematic) role from the verb or not. If they do, they are arguments, and based on our omission test they serve as complements as well. The problem is that the NP arguments of the infinitive get their case, as well as their semantic role from the infinitives alone. This fact does not support the assumption that they are arguments of the matrix verb as well, however it cannot be regarded as a counter-example either.

Examining English raising verbs we find that one of their arguments does not get the thematic role from the verb. These verbs have a predicative complement, e.g. an infinite verb the PRO-subject of which can be coindexed with the complement, therefore getting a thematic role from this infinite verb.

(3) a. *He seems to be old.*
b. *I believe him to be old.*
c. *I believe the desk to be old.*
d. *I believe him to be clever.*
e. **I believe the desk to be clever.*

It can be seen that the constituent *he* gets its case from the finite verb, not the infinite one (sentences (3a) and (3b)). The thematic role, however is given to it by the infinite verb as we can conclude from the opposition between sentence-pairs (3b) and (3c) on the one hand and (3d) and (3e) on the other. Thus, if a regent does not give a thematic role to one of its constituents does not mean that this constituent is not a complement of the verb. So when constructing a grammar we have to decide whether to treat the complements of the infinitive as the complements of the matrix verb or not.

3 Syntactic background

Since a well-motivated decision cannot be merely based on the lexical description of verbs, we have to examine the structure of the sentences containing infinitives. Katalin É. Kiss has already addressed this problem and found that sentences with infinitival constructions show properties of simple and compound sentences at the same time. Let's have a closer look at the phenomena to be explained.

3.1 Simple sentences

On the basis of Hungarian word order, the arguments of the infinitive cannot be separated from those of the matrix verb: they can mix freely with them. Moreover, if the arguments of the infinitive precede the main verb they can function as its focus and topic[3], exactly as if they were the arguments of the finite verb.

---

[3] In Hungarian the focus lies immediately in front of the verb, and it has a stress. Focus-constituents will be printed in capital letters. Topic-constituents lie in the very first part of the sentence, before the focus.

(4) a. *Péter egy könyvet akar Marinak adni.*
P a book-acc wants M-to to-give
'Peter wants to give a book to Mary."
b. *MARINAK akar egy könyvet adni Péter.*
'Peter wants to give a book to MARY.'
c. *Péter ADNI akar Marinak egy könyvet.*
'Peter wants TO GIVE Mary a book.'
d. *AKAR Marinak Péter egy könyvet adni.*
'Peter WANTS to give a book to Mary.'
etc.

In sentences (4abc) it is the infinitive itself (4a) or one of its arguments (4b and 4c) that occupies a position preceding the finite verb. According to a traditional transformational grammar analysis proposed by É. Kiss in a series of papers, these constituents can only get to these positions as a result of transformations. On the basis of É. Kiss' (1989) proposal the first two sentences can be analysed as either a simple or a compound sentence. In the first case the complements of the infinitive move to a position in front of the verb from the neutral argument slots after the finite verb, whereas in the second case the transformation is applied to a constituent of a subclause. This latter phenomenon occurs in (5):

(5) *Péter MARIT akarja, hogy fölvegyék az egyetemre.*
P M-acc wants that they-admit the university-on
'Peter wants MARY to be admitted to university.'

However, sentence (4c) cannot be analysed as a compound sentence, for in this case the verb, that is the head of the subclause should be moved from its position which results in an ungrammatical sentence in case of normal compound sentences.

(6) **Péter FÖLVEGYÉK akarja, hogy Marit az egyetemre.*

There is another reason why sentence (5d) cannot be analysed as a compound sentence: in this example the subject of the main verb (Peter) stands between the arguments of the infinitive (Marinak, egy könyvet). If we want to analyse this sentence as a compound we would either have to assume a downward transformation moving the subject of the finite verb to the subclause, or, as an alternative possibility, when generating the matrix sentence we have to generate an empty position after the verb where a constituent of the subclause can be moved later. Neither of these solutions are really desirable. We have to generate the complements of the infinitive in the complement slots of the finite verb, considering them the complements of this verb.

3.2 Compound sentences

Sometimes the relationship of an infinitive and its arguments is parallel to a traditional view of the relationship between a sentential head and the arguments it selects. These are the relationships between a finite verb and its focus, topic or modifier. Let's see some examples:

(7) a. *Péter szeret el-látogatni Marihoz.*
P likes away-to-visit M-to
'Peter likes visiting to Mary.'

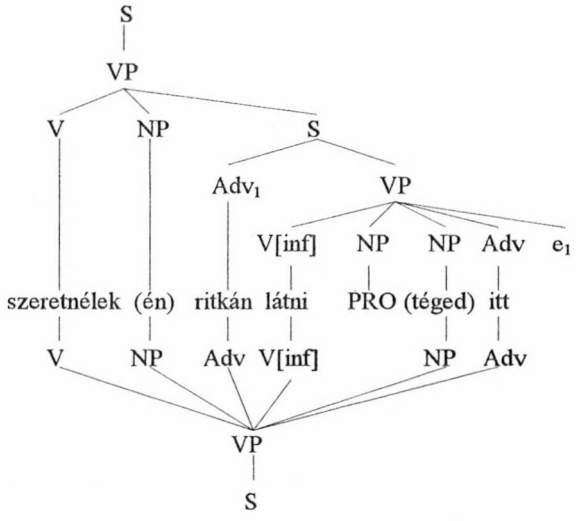
b. *Péter szeretne MARIVAL beszélni.*
P would-like M-with to-speak
'Peter would like to speak with MARY.'
c. *Péter szeretne NEM BESZÉLNI Marival.*
P would-like not-to-speak M-with
'Peter would like not to speak with Mary.'
d. *Péter szeretne NEM MARIVAL beszélni.*
P would-like not-M-with to-speak
'Peter would like to speak not with Mary (but with someone else).'
e. *Péter szeretne mindent megbeszélni Marival.*
P would-like everything-acc PFX-to-speak with M
'Peter would like to discuss everything with mary.'

(7) shows that there can be different constituents that precede the infinitive: verb modifier (7a), focus (7b), and universal quantifier (7e). Constituent negation can occur in the infinitival clause (7d), just like sentential negation (7c). On the basis of these data we can say that the infinitival expression functions as a clause, that is, the sentence containing this infinitival expression is a compound sentence. But then the arguments of the infinitive cannot be the arguments of the finite verb.

3.3 É. Kiss' solution

Earlier we raised the question whether the complement of infinitives are to be treated as belonging to the main verb or not. Now it seems that we have to do both at the same time, simultaneously analysing sentences containing an infinitive as simple AND compound sentences. That is exactly what É. Kiss proposes in her article, namely assigning two structural descriptions to the sentences in question:

(8)



Szeretnélek (én) ritkán látni (téged) itt.
would-like-I-you$_{acc}$ I rarely to-see (you-acc) here
'I would like to see you here rarely.'

At the top of the figure the structure of the compound sentence can be seen, at the bottom the simple one. The phenomena discussed above can be accounted for with the application of both sentence structures.

The problem with this double analysis is that there are grammatical Hungarian infinitival sentences that cannot be properly analysed in either way:

(9) *Ma még szeretne CSAK A KÖNYVRŐL beszélni Péter Marival.*
today would-like only the book-about to speak P M-with
'Peter would like to speak with Mary only about the book today.'

This sentence could not be analysed as a compound sentence within the framework described by É. Kiss, for the subject of the finite verb is in the infinitival expression. However, it cannot be analysed as a simple sentence either, as there is a focus phrase in it (*csak a könyvről*), which could get into a position after the main verb only if there were another focus-constituent in front of the verb:

(10) *PÉTER látta meg csak a könyvet.*
P saw PFX only the book
'It was Peter who saw only the book.'
**Péter meglátta csak a könyvet.*
'Peter saw only the book.'

So the problem is not solved yet.

4 An alternative proposal

My proposal is to change our view of focus. So far we called a constituent focus if it occupied the focus position, which we regarded as part of the sentence structure. That is, we identified the focus with the focus position: if we found a focus in the infinitival expression we had to conclude that this infinitival expression has a focus position, which is an essential part of Hungarian sentence structure.

This problem is easier to solve if we discard the strictly positional analysis of focus. Instead of bounding focus to the sentence structure, let us try to describe it as a relationship between the verb and one of its arguments. According to my proposal it is the regent that can make a focus of one of its complements, that is, every focus is *the focus OF some other constituent*. Put differently, a verb can give a focus feature to one of its argument.

If we accept that the phenomena discussed above can be regarded as the mutual relationship of two constituents, there can be no objection to treating sentences with an infinitival expression as simple sentences. Moreover, the phenomena illustrated with the examples (4) force us to choose this solution. If we analyse them as simple sentences, however, the complements of the infinitive have to be regarded as arguments, therefore complements of the main verb, as the omission test shows. In this way we can give a simple solution to the phenomena discussed in this paper: the mixed occurance of the arguments of the infinitive and those of the main verb is a natural phenomenon, as each of them is a complement of the finite verb. For the same reason is it possible that the infinitive and its complements can function as topic or focus of the main verb. The infinitive itself can also have focus of its own (and also quantifiers and negated arguments), for it still functions as a regent, that is, it can mark one of its arguments as focus.

5 Formal description

How exactly the "kidnapping" of these complements happens can be described within the framework of the Head-driven Phrase Structure Grammar (HPSG) (Pollard, Sag, 1994). This framework is especially good for this purpose, for the complement structures are detailed enough to serve as a basis for each level of the sentence analysis.

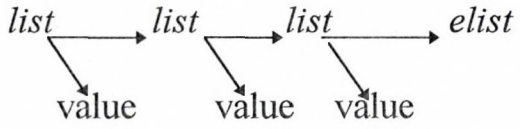
5.1 A brief survey of HPSG

In HPSG the description of linguistic objects is done with the help of *feature structures* independently of whether the unit to be described is lexical or phrasal. Feature structures consist of feature-value pairs. The value of a feature can be another feature structure or an atomic value. Every feature structure is a directed acyclic graph, where features are represented by edges. The value of a feature is the node where the edge points. If a node is the starting point of an edge it represents a feature structure, otherwise it gets an atomic value. The analysis of a sentence is itself a feature structure, built up from feature structures representing the constituens, whose smallest parts are feature structures of lexical units.

The difference between lexical and phrasal units is that only the latter can have the DTRS feature, whose value is a list of feature structures. The elements of this list can be phrasal or lexical units.

An empty list is represented by the atomic value *elist*. Otherwise a list contains two features: one of them is a feature structure or an atomic value representing one element of the list, the other is another list (which is actually the rest of the original list). The graph of a list is shown in (11) (without the names of the features).

(11)



The feature structure of a linguistic unit (a phrase or a lexical unit) contains the phonological description as the value of the PHON feature, various syntactic and semantic information as the value of the SYNSEM feature and, in the case of non-lexical units a DTRS feature, whose value is the list of constituents.

Within the SYNSEM feature we differentiate three further features: CATEGORY, CONTENT and CONTEXT. The value of the CATEGORY feature is a feature structure that contains mainly syntactic information. For instance the HEAD feature gets the values that are shared with the mother node. The detailed description of the HEAD feature structure can be omitted here.

The *valence features* list the complements of a constituent, and they also belong to the CATEGORY feature. In English the valence features are the SUBJ, standing for subject, and a COMPS feature with a list value containing all the complements of the verb, but the subject, in order of obliqueness (the first element being the less oblique). The valence features have the same type of values as the SYNSEM feature.

The value of the CONTENT feature contains the agreement features (INDEX) and the situation semantic description of the object. The values of the CONTENT and CONTEXT features are of no importance for the purposes of the present paper, therefore I omit their detailed description.

For example the feature structure of the English verb *gives* is the following (12):

(12)

$$\begin{bmatrix} \text{PHON: } \textit{gives} \\ \text{SYNSEM:} \begin{bmatrix} \text{CAT:} \begin{bmatrix} \text{HEAD: V[fin]} \\ \text{SUBJ: NP[nom]}_{(1)[3rd,sing]} \\ \text{COMPS: } \langle \text{NP[acc]}_{(2)}, \text{NP[acc]}_{(3)} \rangle \end{bmatrix} \\ \text{CONTENT:} \begin{bmatrix} \text{RELN: give} \\ \text{GIVER: (1)} \\ \text{GIVEN: (2)} \\ \text{GIFT: (3)} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Since *gives* is a lexical unit, it has no DTRS feature.

NP and V are abbreviations describing the matching feature structures.

The right-hand subscript of NP stands for the INDEX feature of the noun phrase, meaning that it has to be third person singular.

The numbers in brackets (tags) mean *structure-sharing*, that is, feature structures marked with the same numbers have the same features with exactly the same values after the tags. In (12) this means that we have to identify the subject of the verb with the semantic role of the *giver*. Due to the structure-sharing there will be acyclical loops in the directed graphs of a feature structure.

Within HPSG rewriting rules are divided into ID (Immediate Dominance) and LP (Linear Precedence) rules.

What is especially appealing in HPSG from the point of view of the present problem is its explicit handling of the set of complements. Due to this, rules and restrictions are easy to formalise, and any change in the complement structure is directly reflected in the sentence structure. A further advantage is that the order of constituents is defined with binary relations, which determine the position of two constituents relative to each other.

5.2 The lexical representation of raising verbs

At first sight it might seem to be a disadvantage that the complement structure is set and detailed in the lexicon, but if we adopt the method used in the treatment of English raising verbs no problem arises. Raising verbs do not have any restriction on their subject or object complement in the lexicon apart from defining its case and prescribing that it has to be structure-shared with the subject of their predicative complement. This predicative complement does not provide its subject with case, as non-finite verbs do not define the case of their subjects (see (3)) (like in GB).

The description of the SYNSEM feature of the verbs *seem* and *believe* is as follows:

(13) a.

$$\begin{bmatrix} \text{CAT:} \begin{bmatrix} \text{SUBJ:(2)[nom]} \\ \text{COMPS:} \langle \text{VP[inf, SUBJ:(2)]:(1)} \rangle \end{bmatrix} \\ \text{CONTENT:} \begin{bmatrix} \text{RELATION: seem} \\ \text{SOA - ARG: (1)} \end{bmatrix} \end{bmatrix}$$

b.

$$\begin{bmatrix} \text{CAT:} \begin{bmatrix} \text{SUBJ: NP[nom]}_{(1)} \\ \text{COMPS: } \langle (2)[\text{acc}], \text{VP[inf, SUBJ: (2)]:(3)} \rangle \end{bmatrix} \\ \text{CONTENT:} \begin{bmatrix} \text{RELATION: believe} \\ \text{BELIEVER: (1)} \\ \text{SOA – ARG: (3)} \end{bmatrix} \end{bmatrix}$$

As we can see, the complement of the verb bearing tag (2) is the same as the subject of the infinitive. This structure-sharing does not result in a cyclic graph.

The tags on the right side of the COMPS list ((1) and (3) respectively) refer to the value of the CONTENT feature of the verb phrase. In (13b) this means that the thing believed is the state of affairs the verb phrase expresses.

5.3 Application

Hungarian verbs with infinitival complements are related to English raising verbs so far as they treat one or more complement of their infinitives as their own complements. The English verb provides them with case, the Hungarian with focus feature, but none provide them with a semantic role. The difference is that while raising verbs do this with only one specified complement of the infinitive, their Hungarian counterparts do it with all the complements of the infinitive but the subject.

When adopting the treatment described above we must not forget that we may not know exactly how many arguments the infinitive has. The nature of lists (see (11)) can be a solution to this problem. By definition we do not know the exact number of the elements in a non-empty list, all we know about it is that it has at least one element. The exact number of the elements is irrelevant to the purposes of the description. So the COMPS list of verbs like *szeret* will include its own complements and in addition, the complements of its infinitive as well:

(14)

$$\begin{bmatrix} \text{PHON: } \textit{szeret} \\ \text{SYNSEM:} \begin{bmatrix} \text{CAT:} \begin{bmatrix} \text{HEAD: V[fin]} \\ \text{SUBJ: NP}_{(1)[3\text{rd,sing}]}\text{[nom]} \\ \text{COMPS: } \langle \text{V[inf, SUBJ: NP}_{(1)}\text{, COMPS:(2)]:(3)} \rangle + (2) \end{bmatrix} \\ \text{CONTENT:} \begin{bmatrix} \text{RELATION: like} \\ \text{LIKER: (1)} \\ \text{SOA - ARG: (3)} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

The symbol + stands for the concatenation of two lists. That (2) is really a list, not only an element of the list can be seen from the fact that it is introduced as COMPS: (2), which means a list of complements as opposed to COMPS: <(2)> which stands for one element of a given list of complements. Such a concatenation of elements is allowed in HPSG, it does not lead to inconsistency within the system, for the graph representing the resulting feature structure does not contain a cyclic loop.

6 Verb carriers

I have already mentioned and it can be clearly seen from the feature structure (14) that verbs with infinitival complements have a special relationship to the subjects of their infinitives. While the rest of the complements are simply copied to the complement list of the matrix verb, the subject of the infinitive is merely co-indexed with the subject of the matrix verb. It can be stated in general that the subject complement of a verb is different from the rest of the complements. This made the founders of HPSG divide the subcategorization frame of the verb into the features of SUBJ and COMPS (9th section in Pollard-Sag, 1994). In the case of Hungarian, however, it is not only the subject that is a separate unit, but another type of complements as well: the verb modifier, or, as Kálmán et al. call it, the verb carrier (Kálmán et al., 1989).

Verb carriers are given a special treatment: they are the closest relations of the verb, and as such when the sentence is neutral, they stand immediately on the left side of the verb. The verb carrier forms an independent phonological and syntactic unit, but from a semantic point of view it cannot be separated from the verb, especially if we think of the non-compositional carrier-verb expressions such as *becsap* (in-strike, 'make a fool of'). When a verb carrier is part of a sentence structure containing an infinitive, its behaviour is particularly awesome.

A group of the verbs with infinitives require a verb carrier (these verbs are called auxiliaries by Kálmán et al. 1989). In case the infinitive does not have an own verb carrier, the infinitive itself functions as the verb carrier of the finite verb (see sentences (15ab)). Otherwise the verb carrier of the infinitive is the verb carrier of the finite verb as well (15cd):

(15) a. *Péter úszni akar.*
P to-swim wants
'Peter wants to swim.'
b. *PÉTER akar úszni.*
P wants to-swim
'It is Peter who wants to swim.'
c. *Péter el akar menni.*
P away wants to-go
'Peter wants to leave.'
d. *PÉTER akar el-menni.*
P wants away-to-go
'It is Peter who wants to leave.'

Bearing these in mind, I put forward the following proposal: in the same way as the subject was treated, we should separate the verb carriers from the rest of the complements. So a new valence feature is to be introduced at the lexical description of a verb, that of CARRIER, the value of which is the carrier of a given verb, or, if it does not have one, the *empty* value. Of course different lexical rules may change this; e.g. if a verb with a verb carrier is turned into a verb with focus (that is the, focus feature appears on the verb and on one of its complements) its carrier feature is deleted, the value of it appears on the COMPS list, that is, the earlier carrier turns into an ordinary complement.

Therefore, verbs with infinitives that do not have a verb carrier have the structure described under (14) with a slight modification: the feature-value pair of CARRIER: *empty* is added to them. The valence features of the auxiliary verbs will be formalised in the following way:

(16) a.

$$\begin{bmatrix} \text{SUBJ: NP[nom]}_{(1)} \\ \text{COMPS: } \langle\ \rangle + (2) \\ \text{CARRIER: V[inf, SUBJ: NP}_{(1)}\text{, COMPS: (2), CARRIER: empty]} \end{bmatrix}$$

b.

$$\begin{bmatrix} \text{SUBJ: NP[nom]}_{(1)} \\ \text{COMPS: } \langle \text{V[inf, SUBJ: NP}_{(1)}\text{, COMPS: (2), CARRIER: (3)]} \rangle + (2) \\ \text{CARRIER: (3)} \end{bmatrix} \wedge (3) \neq \text{empty}$$

7 Conclusion

The subject of the present paper is the description of the relationship between verbs with infinitives and the complements of the infinitives they take. Based on the reasoning described above we can conclude that these complements are the complements of both the infinitive and the matrix verb. The lexical representation of the verbs in question is carried out in line with the treatment of raising verbs in HPSG.

Presumably the proposal put forward can be applied not only to the treatment of infinitival sentences but other similar constructions; for instance sentences containing copulas. Hence the expression in the title 'complement of complements' rather than 'complements of infinitives'. Let's see an example:

Tegnap heves vita volt a parlamentben a költségvetésről.
yesterday heated debate was the Parliament-in the budget-about
'There was a heated debate about the budget in the Parliament.'

In this sentence the matrix verb (*volt)* treats the complement (*a költségvetésről)* of its nominal complement (*vita)* as its own. So the paper does not merely present an alternative solution for the treatment of infinitival sentences but introduces a device that can be fit for application in further areas.

References

Kálmán C., György - Kálmán, László - Nádasdy, Ádám - Prószéky, Gábor. 1989. A magyar segédigék rendszere. *Általános Nyelvészeti Tanulmányok* XVII:49-103. Akadémiai Kiadó, Budapest

Komlósy, András. 1992. Régensek és vonzatok. In *Strukturális magyar nyelvtan*, ed. Kiefer Ferenc. Akadémiai Kiadó, Budapest

É. Kiss, Katalin. 1989. Egy főnévi igeneves szerkezetről. *Általános Nyelvészeti Tanulmányok* XVII:153-169. Akadémiai Kiadó, Budapest

Pollard, Carl, and Sag, Ivan A. 1994. *Head-Driven Phrase Structure Grammar.* Chicago & London, The University of Chicago Press.

Department of General Linguistics
József Attila University
Egyetem u. 2. H-6722 Szeged, Hungary
e-mail: altnyelv@primus.arts.u-szeged.hu

On affricates*

Péter Szigetvári

The representation of affricates (and other "two-phase" segments) is far from being undisputed in phonological theory. Some researchers have even questioned whether affricates are monosegmental or rather consonant clusters. Though at present there seems to be consensus on the monosegmental analysis, how this segment should be thought of is a debated issue. This paper collects evidence for and against various views on affricates and it will offer a solution in a unary-element framework which dispenses with any meaningful distinction between contour and non-contour segments. As this paper is planned to be the outset of a larger scale work on segmental representations, I will leave a large number of problems open for further research.

A typical definition of an affricate runs as "a stop released into the homorganic fricative within one and the same syllable and one and the same morpheme" (Catford 1988 : 112). The first part of this definition belongs to the domain of phonetics, while the second part is phonological: syllables are theoretical constructs, the location of a syllable boundary is very often theory-dependent, and morpheme boundaries are even more remote from the physical properties of the speech signal.

A stop followed by a homorganic fricative may phonetically be treated as an affricate, such is often the case with sequences like *ca*[ts], [tr]*ue*. This, however, has little bearing on the issue whether this physical event is one phonological unit or two, that is, whether in an autosegmental representation it should occupy one slot on the timing tier or it should spread out holding on to two. Although the postvocalic parts of *cats* and *catch* sound similar (apart, of course, from the place of articulation), and the initial parts of *true* and *chew* are even closer to each other physically, the relevant portions of the first members of these pairs are almost unanimously considered two units by phonologists, those of the second members are treated as monosegmental: affricates. In this paper I am primarily concerned with the phonological representation of affricates, and will, therefore, disregard such phonetic similarities.

1 Affricates as consonant clusters

Looking for reasons for considering English [ʧ] and [ʤ] consonant clusters one may come up with the fact that these two sound (sequence)s do not pattern with some other, doubtlessly monosegmental sounds, like [t k p f], in branching onsets: beside the possible word-initial onset clusters [tr kw pl fj], *[ʧl], *[ʤw] and the like are ungrammatical. This fact would be neatly explained by supposing that affricates are bisegmental, thus a binary branching onset is saturated by an affricate. In this case *[ʧl] is just as impossible as *[plj].[1]

Another piece of argument for not treating affricates as monosegmental is that they are excluded from "genuine", i.e. word-medial, coda positions, at least in English.[2] The affricates [ʧ] and [ʤ] are either followed by a vowel or by a word- (strong morpheme-) boundary, #. It is not only the two affricates, however, that cannot occupy non-word-final

* I am grateful to Péter Siptár for pointing out a number of typos, stylistic and grammar mistakes and especially for the idea of ranking constraints on head-deletion.

[1] Interestingly, *#[stʃ] is also unprecedented. This fact, however, does not corroborate the bisegmental view, since *s*-initial clusters can contain up to three consonants in English (cf. [str spl]), so even if the affricate were a consonant cluster like [tr] this would not immediately exclude it from this position. Instead, we may suspect an accidental gap here: even word-medially [stʃ] occurs only in a handful of words.

[2] The codahood of word-final consonants has been seriously threatened (cf. Kaye 1990); in many languages word-final consonants share at least as many properties with onsets as with codas.

coda position in English but also two stops, [t] and [d].[3] This means that the restriction on codas is more general: there is a ban on any [+cor, −cont, −son] segment, i.e. [t d ʧ ʤ], in this position, which then has nothing to do with the mono- or bisegmental status of the consonants in question.

The branching-onset argument exposed above can also be partly explained away. A weak universal restriction on branching onsets is that the segments occupying the two positions cannot be homorganic.[4] Thus *[pw], *[tl], *[ʃj] are ruled out, and this also excludes *[ʧr] and *[ʧj].[5] *Cw* clusters are rather rare in English anyway; for the non-occurrence of *[ʧl] and *[ʤl] one may blame a historical conspiracy: there was no input in earlier stages of the language that any change could have turned into these clusters.[6] Besides these alternative reasons for affricates not turning up in branching onsets in English, it is not even true that an affricate cannot be the first segment in an onset. German provides examples like *Pflanze* 'plant', *Pfriem* 'awl', *zwanzig* [t͡sv] 'twenty'.[7]

Japanese and Hungarian also provide evidence that the consonant cluster view is untenable. Both languages "dislike" branching onsets: Japanese does not have word-initial consonant clusters at all, while Hungarian has them only in words that are usually felt to be foreign by native speakers (Törkenczy 1989). Despite these facts both languages have word-initial affricates: J [t͡suki] 'moon', H *cica* [t͡sit͡sɒ] 'kitty'. Even in languages that do have branching onsets without any restrictions affricates would exemplify the only stop+fricative clusters in this position. English does not feature branching onsets word-finally, hence *cycle* *[-kl] → [-kəl], affricates, however, are not affected by this restriction, they freely turn up at the end of words. Polish definitely forces the analyst to distinguish affricates and stop+fricative clusters, since it makes phonological use of this distinction in pairs like *czy* [ʧɯ] 'whether' vs. *trzy* [tʃɯ] 'three' or *czech* [ʧex] 'Czech' vs. *trzech* [tʃex] 'three GEN' (Jakobson *et al.* 1952). To add a last argument against the cluster view one might mention English stress assignment which treats both parts of an affricate as extrametrical: *mána⟨ge⟩*, for example, patterns like *édi⟨t⟩* and unlike *tormén⟨t⟩* (cf. Harris 1994:40).

We have seen that the overwhelming majority of arguments support the monosegmental analysis of affricates. What the representation of this one segment should be is a much less settled issue. This is what we are going to turn to presently.

2 Data and views on affricates

The different views on the status of affricates may be classified into two groups: for the better part of this century they were thought of — if monosegmental — as some (special) kind of plosives (e.g. Jones 1972:158; Jakobson *et al.* 1952; Laziczius 1963:61; SPE:177 and 319), while from the 1970s a set of alternative analyses has evolved, which all share the idea of assigning two contradictory values to the feature [continuant] (Hoard 1971; Halle & Clements 1983; Sagey 1986; Hualde 1988; Lombardi 1990). Some researchers have lately returned to the affricates-are-plosives line of research (Steriade 1992, 1993, 1994; Rubach 1994; Schafer 1995), which this paper is also going to argue for.

[3] Word-internal *t/d+C* clusters in English either contain a word boundary (see last note), or are the result of syncope, in which case the stop may be assumed to be in onset position followed by an unpronounced nucleus.

[4] The constraint is violated in, for example, Polish: *błoto* [bw-] 'mud', *pług* [pw-] 'plough'.

[5] For gauging onset homorganicity, [r] must be treated as palatal to allow for [tr]. The cluster [ʃr] seems to contradict the restriction, but there is reason to suppose that it does not constitute a branching onset: its status is similar to *sC* clusters, whatever their status may be (cf. Kaye 1992).

[6] Alternatively, [ʧ ʤ] (as well as [ʃ ʒ]) may be classified as both coronal and palatal, which renders *[ʧl] ungrammatical, but one would still need the "conspiracy" escape hatch for *[ʧw].

[7] Here again the ban on homorganicity is respected, hence *[p͡fv], *[t͡sr], *[t͡sl].

2.1 Pre-autosegmental proposals

The Jakobsonian idea of affricates being strident plosives (Jakobson *et al.* 1952) has been put aside by Chomsky & Halle in their revised feature theory on the grounds that there exist non-strident affricates which do in certain languages (e.g. Chipewyan) contrast with strident ones (SPE: 321–322). What the SPE proposes is the distinction instantaneous release vs. delayed release, where the former characterizes "normal" plosives and the latter affricates. This renders the specification of stridency rather redundant: it now distinguishes affricates in, say, Chipewyan and also certain pairs of fricatives, but these latter also differ in their place of articulation usually. What's more, in this new framework [+strident] is incompatible with [−continuant, −delayed release]. Also, the very natural change $t+s \rightarrow \widehat{ts}(:)$ becomes very unnatural to express, as shown in (1):

(1) **$t+s \rightarrow \widehat{ts}$ as a feature-matrix changing rule**
(place and laryngeal specifications omitted)

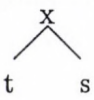
$$\begin{bmatrix} -\text{son} \\ -\text{cont} \\ -\text{del rel} \\ -\text{strid} \\ \vdots \end{bmatrix} + \begin{bmatrix} -\text{son} \\ +\text{cont} \\ -\text{del rel}^{8} \\ +\text{strid} \\ \vdots \end{bmatrix} \longrightarrow \begin{bmatrix} -\text{son} \\ -\text{cont} \\ +\text{del rel} \\ +\text{strid} \\ \vdots \end{bmatrix}$$

Its narrow range of applicability called for alternative models to replace the feature [del rel] very soon after its introduction. Hoard (1971) in his review of the SPE points to two other phenomena that disfavour the [del rel] representation of affricates: (i) the rule $s+t \rightarrow \widehat{ts}(:)$ is just as marked as its oposite (above), although its occurrence Hoard claims to be very unlikely[9] and (ii) Quileute (an Amerindian language spoken in British Columbia) features the rule $tʃ \rightarrow ʃ \,/\, k\_\_$, which again needs arbitrary feature value changing rules instead of the intuitively obvious deletion of the stop phase of the affricate. What Hoard proposes in order to dispense with this feature is a feature matrix that contains two, sequentially fixed, instances of the feature [continuant], the first with a −, the second with a + value. This, however, goes totally against the type of feature matrices Chomsky & Halle propose.

2.2 The orthodox contour segment analysis

Hoard's idea was revived in autosegmental frameworks, which let the analyst posit one-to-many relationships between features and time segments they are associated with. A typical representation thus looks like the one in (2):

(2) **The representation of [t͡s] in traditional autosegmental phonology**

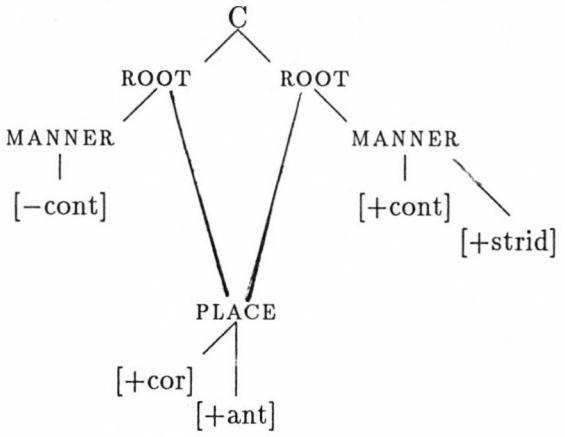


There are several problems with this representation. If the letters are taken, as is usual, to stand for abbreviations of feature matrices, several features (the place, source and major class features) are given twice, which is redundant. It is arbitrary then that there should not exist affricates whose two phases had different place of articulation or differed in voicing. Invoking some type of feature geometry, the model may be simplified having independent features hanging only from the Manner node (Polgárdi 1991: 13):

[8] Some phonologists characterize fricatives and also some liquids as [+del rel] (cf. Siptár 1994: 214). This unique move makes the change more natural.

[9] My own observations of child language do not support Hoard's claim: $s+t \rightarrow \widehat{ts}(:)$ does occur.

(3) **Modified representation of [t͡s]**



(Due to representational difficulties, the laryngeal specifications, which are also shared by the two root nodes, are omitted from the diagram.) The fact that it is the Manner node which holds the features that may be distinct in the two phases is felicitous: in the two usual types of contour segments, affricates and prenasalized stops, the difference between the first and the second half is in the values of [continuant] and [strident] in affricates and in that of [nasal] in prenasalized stops, all of which are dependent on the Manner node.

The underlyingly fixed order of the two halves of an affricate is supported by a set of so-called edge phenomena. In processes whose trigger is to the left affricates often pattern with stops, while in processes that see affricates from the right they behave similarly to fricatives. Zoque (Penutian; Mexico) provides us with an example: postnasal stops are voiced (e.g., /min-pa/→[minba] 'he comes'), whereas fricatives in the same position are not influenced (e.g., [winsaʔu] 'he received'). Affricates are voiced by a preceding nasal, just like stops (e.g., /pʌn-ʧʌki/→[pʌn-ʤʌki] 'figure of a man'; examples from Sagey 1986 quoted in Kenstowicz 1994 : 500). Preglottalization, typical of post-tonic voiceless stops in English, affects the voiceless affricate [ʧ] as well. English plural and past allomorphy shows that affricates pattern with fricatives from the right: they select the vowelless [t]/[d] past allomorph, but the vowelful [ɪz] for the plural, which is what [s z ʃ ʒ] and the opposite of what [t d] do.

Affrication processes like *t+s→t͡s(ː)* also call for an ordered contour segment analysis, derivable from some kind of merger of the adjacent stop and fricative. The same mechanism, however, cannot be applied to other frequently occurring changes resulting in affricates, e.g., *t+j→ʧ(ː)* in English and Hungarian.

2.3 The unordered-feature contour segment analysis

Besides edge phenomena we also encounter facts that argue against a fixed sequential order in contour segments, more specifically, affricates. Basque is particularly rich in such anti-edge phenomena, some of which are illustrated below based on Hualde (1988).

Basque stops are voiced when preceded by a nasal or a lateral:[10] /neka-tu/ → [nekatu] 'get tired PERF', /ar-tu/ → [artu] 'take PERF', but /afal-tu/ → [afaldu] 'have dinner PERF', /ken-tu/ → [kendu] 'take away PERF'. Contrary to expectations affricates fail to be influenced by a preceding nasal or lateral: /neka-tsen/ → [nekatsen] 'get tired IMPERF', /ar-tsen/ → [artsen] 'take IMPERF', /afal-tsen/ → [afaltsen] 'have dinner IMPERF', /ken-tsen/ → [kentsen] 'take away IMPERF'.[11] Fricatives are not voiced either,

[10] I am using Hualde's transcription of the Basque examples.

[11] It is true that most Basque dialects do not have /d͡z/ in their inventory, but even those which do leave the affricate voiceless.

but they are strengthened to affricates in this position: /mendi-sale/ → [mendisale] 'mountaineer', but /aran-sale/ → [arantsale] 'fisherman'.

Another rule palatalizes coronal noncontinuants that occur after a high front vowel or glide: /itaun/ → [it'aun] 'question', /neska-tila/ → [neskatiʎa] 'girl', /ipin-i/ → [ipiñi] 'put PERF'. Affricates, as well as fricatives, are exempt from this rule: /itsar-i/ → [itsari] 'awake PERF', /isen/ → [isen] 'name', despite the fact that [ʧ] and [ʃ] both occur in Basque.

Whereas in the two preceding rules affricates did not undergo a change they were expected to, in the following cluster simplification process the affricate is affected by the rule although it should not be. Basque stop+occlusive[12] clusters are simplified by losing the first stop: /bat paratu/ → [baparatu] 'put one', /bat-naka/ → [banaka] 'one by one'. Fricatives do not take part in the process, while affricates, in this case located to the left of the trigger and thus showing their fricative face, are nevertheless simplified into a fricative: /its-tegi/ → [isteɣi] 'dictionary', /arits-mendi/ → [arismendi] 'oak mountain'.

Lombardi (1990) cites Yucatec Mayan examples, in which the first element of homorganic stop+plosive sequences turn into [h] (*tʧ→hʧ, kk→hk*), but affricates in the same environment become fricatives (*t͡st→st, ʧt→ʃt*): both stops and affricates lenite, although the latter show their fricative face towards the trigger. Turkish also offers an example of an anti-edge phenomenon (Hualde 1988, Lombardi 1990). Word-final noncontinuant obstruents are devoiced in this language, but fricatives remain voiced: [kanat] 'wing' ([kanadɯ] 'wing ACC') vs. [kɯz] 'daughter'. Affricates pattern with stops in devoicing word-finally: [pabuʧ] 'slipper' ([pabud͡ʒu] 'slipper ACC'). This defies our expectations, since the trigger, the word boundary, is adjacent to the fricative phase of the affricate.

Proponents of unordered contour segments thus retain the [−cont], [+cont] specifications (or, rather, the [stop] and [cont] privative features), but argue that the two features are not ordered lexically and throughout the phonological derivation. It is only the phonetic interpretation that introduces their order, which is predictable, hence redundant underlyingly.

2.4 An edge–anti-edge phenomenon

Hungarian adaffrication exhibits both edge and anti-edge phenomena. The following chart summarizes the changes, the affricate outputs are set in boldface:

(4) **Hungarian adaffrication**

	s	t͡s	t
s	sː	st͡s	st
t͡s	**t͡sː**	**t͡sː**	t͡st*
t	**t͡sː**	**t͡sː**	tː

*[st] is also possible in fast speech

As the first element of a cluster, affricates pattern with stops, despite the fact that it is their fricative phase which is closer to the other segment involved in the change. As the second element, they pattern with fricatives, although this time it is their stop phase that is towards the other component of the change. Preceding a stop, however, an affricate behaves like a fricative: neither forms a geminate with the stop. An attempt at an explanation will be made in section **3.1**, where the process will also be discussed in more detail.

2.5 Rubach's proposal

Summarizing the predictions of the ordered and the unordered contour segment analyses, Rubach (1994) concludes that the former fares better in the Dental Spreading rule of Polish

[12] I am using the term "occlusive" to include stops, affricates and nasals ([−cont]), while by "plosive" I mean stops and affricates ([−cont, −son]).

(in which [s] spreads its place on the preceding sibilant, [t͡s] does not), and also in Nasal Gliding ($n \rightarrow \tilde{w}$, $ɲ \rightarrow \tilde{j}$ before a fricative, but not before an affricate). Even the ordered contour segment analysis fails, however, in the case of Fricative Assimilation: [s] and [z] assume the place of articulation of any following postalveolar or prepalatal consonant, but [t͡s] and [d͡z] fail to do so, unless the following consonant is [ʂ ʐ ɕ ʑ] or their affricate counterparts. This analysis also runs astray in Strident Assimilation, where [s z t͡s d͡z] assimilate to the place of articulation of the following strident consonant and, in addition, [t d] become affricates.

As a solution Rubach offers the revival of the Jakobsonian idea of affricates as stops distinguished by the feature [+strident]. Although the rules he brings up are better explained by this hypothesis than by any of its competitors, the whole device is set in a framework that applies orthodox rewrite rules that are subject to rule ordering, mechanisms that are probably too strong for the description of natural languages.

2.6 Schafer's model

In a recent paper Schafer (1995) proposes an asymmetrical relationship between Lombardi's two privative features [stop] and [cont]. Stops contain the feature [stop], fricatives [cont], both residing on the primary stricture tier. Affricates on the other hand have two stricture tiers, with the feature [stop] on the primary and [cont] on the secondary stricture tier. The [cont] feature of the secondary stricture tier is dominated by the [stop] of the primary stricture tier. Lenition of stops and affricates to fricatives results from the association of a [cont] feature to the segment, but being located on the primary stricture tier this feature substitutes the [stop] feature of the target. This configurational model thus predicts that affricates and stops pattern together, since both have the feature [stop] on the primary stricture tier. Schafer cites data from the Toscana dialect of Italian and the Tsimshian Salish language Nisgha to show the prediction to be correct.

In discussing the postnasal strengthening in Tswana, Schafer's representation successfully demonstrates why affricates do not strengthen to stops, but has some difficulty in accounting for the different behaviour of fricatives.

2.7 Steriade's model

Steriade (1992) observes that: (i) only plosives (stops and affricates) can be contour segments, (ii) and they can be contour segments only if they are released and (iii) contour segments never exceed two articulatory phases. She claims, in effect, that not only affricates and prenasalized stops but any stop which is released is a contour segment, since all of them contain two aperture nodes, which are "rather similar to the feature-geometric notion of root-node; [they have] the same functions of anchoring segmental features like place of articulation, nasality, and laryngeal features, and of connecting segments to prosodic structures such as syllables and moras" (Steriade 1993 : 401).

For consonants Steriade proposes three types of aperture nodes: A_0 representing "total absence of oral airflow" ([−cont]), A_f for a "degree of oral aperture sufficient to produce a turbulent airstream" ([+cont, −son]), and A_{max}, which is the "maximal oral aperture in consonants" ([+cont, +son]) (*op.cit.* : 402). A released stop contains the two aperture nodes A_0A_{max}, while for an affricate these are A_0A_f, a solution somewhat reminiscent of the SPE's [instantaneous release] vs. [delayed release] features. Plosive releases, however, are not present underlyingly, but are projected during the derivation by the following universal process (*op.cit.* : 404):

(5) **Universal projection of plosive releases**
$A_0 \rightarrow A_0A_{max}$

The projection of A_{max} vs. A_f depends on place of articulation or, less frequently, on aspiration, on a language specific basis. Thus a labial stop typically projects A_{max}, while a

labiodental will have an A_f release. Such a solution presupposes that no language contrasts a stop and an affricate at the same place of articulation and with the same laryngeal specification. This position is threatened by the fact that the place of articulation of Hungarian [t] and [t͡s] appear to be indistinguishable by phonological features: both are apical and dentialveolar. Maddieson lists six other such languages: Tuva, Tamang, Tzeltal, Squamish, Standard Thai and Malgasy (1984:207, 221), and Rubach (1994:121) points out the same difficulty Steriade has to face in the case of Polish [s] and [t͡s]. The claim that the manner specification of stops and affricates is not distinct underlyingly thus seems untenable.

3 A unary elemental approach

The representation I am to propose for affricates is couched in a framework that posits unary, independently pronounceable elements as phonological primes (Kaye *et al.* 1990, Harris & Lindsey 1995, Brockhaus 1995). In other words, the building blocks of sounds are other, more basic sounds (called elements), which themselves are atomic. It is standardly assumed that one of the elements is special in being the head of the expression forming a segment. The acoustic (and articulatory) properties of the head are modified by the salient properties of the other elements within the expression. Given two elements like U=[w][13] with labiality as its salient property and **h**=[s] with noise as its salient property, **U.h** (where the head appears after the dot by convention) yields a labial noise, [f], **h.U**, on the other hand, is a noisy labio-velar, [ʍ].

The following chart introduces the elements relevant for the discussion, based on Harris (1990), Brockhaus (1995:105) and Szigetvári (1996):

(6) **Some unary elements**

ʔ = [ʔ]	occlusion, abrupt spectral change
h = [s]	narrowed, noise, stridency
H = [h]	spread glottis, aspiration, rise in pitch
R = [ɾ]/[ɜ]	coronality, rise in spectral amplitude[14]
I = [j]/[i]	palatality, large spectral gap
U = [w]/[u]	labiality, fall in spectral amplitude

These elements can be likened to features of orthodox feature theories: e.g., **ʔ** is similar to [−cont] (or [stop]), **H** is like [+spread glottis] and **R** parallels [+cor]. There is no equivalent of [+cont], [−spread glottis] or [−cor], however. The element **h** represents the noise typically accompanying obstruents: it is present in released stops, affricates and (most) fricatives. If anything then, it can be compared to [−son]. There are, nevertheless, major differences between elements and features. Elements are necessarily privative, we have seen that **ʔ**'s ([−cont]) opposite value, [+cont], cannot be expressed in any way, and the same is true of all other elements. Lombardi's and Schafer's apparently privative [stop] and [cont] features are in fact two denominations for the two values of an equipollent feature. Elements are also different in not being abstract phonological constructs like features, but pronounceable sounds. The element **R**, for example, besides being [+cor] is also [+cont], [+son], etc., it is [ɾ] (when consonantal or [ɜ] when vocalic). Properties internal to the sounds that elements encode are inaccessible to the phonological machinery.

Using the elements introduced in (6) we may posit the segmental representations displayed in (7):

[13] When dependent on a nuclear slot, the interpretation of **U** is [u].

[14] Two comments are due here: (i) the existence of this element is debated, I will, nevertheless, make reference to it for simplicity's sake and since here nothing crucial hinges on it and (ii) the two interpretations depend again on whether a non-nuclear or nuclear slot dominates it.

(7) a. [t]=**Rh.ʔ**
b. [t^{h}]=**RHh.ʔ**
c. [t$^{\urcorner}$]=**R.ʔ**
d. [t͡s]=**R?.h**

The representation of the coronal stop, [t], is **Rh.ʔ**, that is, a coronal noisy occlusion (7a). An aspirated stop includes the element **H** in addition to those yielding its unaspirated version: [t^{h}] can be represented as **RHh.ʔ** (7b) (cf. Kaye *et al.* 1990 : 216; Harris 1994 : 133ff.). What an unreleased stop lacks is the noisy burst that accompanies the release phase. Since **h** is responsible for this noise, [t$^{\urcorner}$] is best modelled as **R.ʔ** (7c) (cf. Harris 1990 : 280). The proposal being made here is that by promoting the noisy **h** element to head position in the segment we obtain a plosive which is different from a normally released plosive by being more noisy, that is, an affricate.

The fact that plosives in word-final consonantal positions may and in word-internal codas must lose (or, rather, simply lack) the **h** element may be an instance of lenition, which is typical of these positions. Hungarian, as we will see in the next section, provides a worrying counterexample to this generalization. It is also intriguing why a stop in onset position followed by a pronounced vowel must contain **h**. Steriade, who also encodes the phonologically nondistinctive difference between released and unreleased stops in the representation, proposes a universal release projection rule (5). In our case this would amount to introducing **h** from outside the representation in certain prosodic configurations, certainly an undesirable development. So we have to content ourselves with accepting the impossibility of an **h**-less stop in prevocalic position.[15]

3.1 Some analyses

One of the arguments in favour of the ordered contour segment analysis was the frequent occurrence of the affrication process: *t+s→t͡s(ː).* This change is part of a set of changes dubbed adaffrication and described by Polgárdi (1991) and Siptár (1994 : 210). The changes are listed in (8):[16]

(8) **Adaffrication**
a. t+s→t͡sː, e.g., *ötször* 'five times', *ötödször* 'for the fifth time'
a′. t+ʃ→ʧː, e.g., *barátság* 'friendship', *szabadság* 'freedom'
a″ c+s→t͡sː, e.g., *füttyszó* 'whistle sound', *négyszer* 'four times'
a‴ c+ʃ→ʧː, e.g., *agysejt* 'brain cell'
b. t+t͡s→t͡sː, e.g., *hat cica* 'six kitties', *vad cica* 'wild kitty'
b′. t+ʧ→ʧː, e.g., *hat csók* 'six kisses', *vad csók* 'wild kiss'
b″ c+t͡s→t͡sː, e.g., *nagy cézár* 'great Caesar'
b‴ c+ʧ→ʧː, e.g., *nagycsütörtök* 'the day before Good Friday'
c. t͡s+s→t͡sː, e.g., *malacszerű* 'pig-like'
c′. t͡s+ʃ→ʧː, e.g., *akácsor* 'row of acacias'
c″ ʧ+s→t͡sː, e.g., *csecsszopó* 'suckling', *bridzsszék* 'chair for playing bridge'
d. t͡s+t͡s→t͡sː, e.g., *malaccomb* '(pig's) ham'
d′. t͡s+ʧ→ʧː, e.g., *malaccsülök* 'pig's hoof'
d″ ʧ+t͡s→t͡sː, e.g., *ácsceruza* 'carpenter's pencil'

[15] If the [t] of a [tr] cluster could be proven not to contain h, the nonexistence of affricate+liquid/glide clusters would gain an explanation: in lack of h a stop and an affricate cannot be distinguished. Steriade proposes that such [t]'s only have an A_0 aperture node as a result of release merger, that is, they have no release (1994).

[16] Hungarian spelling quite faithfully renders underlying consonantal segments. The idiosyncratic letter-to-sound correspondences involved are the following: *sz*=[s], *s*=[ʃ], *c*=[t͡s], *cs*=[ʧ], *dzs*=[ʤ(ː)], *ty*=[c], *gy*=[ɟ].

I have ignored the laryngeal distinctions in the changes because their inclusion would have made the number of changes almost four times as many without much use: the laryngeal specification of the resulting affricate is always that of the second component of the input. In the examples both voiceless and voiced cases were given when possible. All the processes in (8) result in geminate affricates.[17] The inputs in (8a) are stop+fricative clusters, in (8b) stop+affricate clusters, in (8c) affricate+fricative clusters and in (8d) affricate+affricate clusters. The place of articulation is determined by the second component, similarly to the case of laryngeal properties.

In formulating rules for these changes in the ordered contour segment framework Siptár (1994 : 260ff.) runs into several problems. He has a uniform rule dealing with stop+fricative and the stop+affricate affrication (which could only be done by two SPE-type rules), but then has to formulate a separate rule for stop+affricate affrication, which occurs in itself, too, as (8b) demonstrates. The unary element model copes with all four changes under the influence of two rather universal constraints, as proposed in (9):

(9) **Coda conditions**[18]
   a. Coda cannot independently license ʔ.
   b. Coda cannot simultaneously license **h** and ʔ.

Condition (9a) states that a noncontinuant segment, which contains ʔ, cannot be followed by a continuant, which does not contain ʔ. Thus in the optimal coda–onset cluster the onset licenses a ʔ element, either on its own or shared with the preceding coda. A situation violating the condition typically arises in nasal+fricative clusters, and can be cured in one of two ways: (i) by postnasal hardening, the coda containing the nasal passes the burden of having to license the ʔ to the following onset, thus making it an affricate or a stop (cf. English *prin*[t]*ce*, also see Steriade 1993 : 420ff.) or (ii) by vocalization, deleting the ʔ, which may result in the total loss of the nasal consonant and possibly leaving residual nasalization on the preceding vowel (cf. Polish *kunszt* [kuw̃ʃt] 'art', Hungarian *kunszt* [kũst] 'trick', Lithuanian *sąžinė* [saːʒineː] 'conscience', which contains a *saN-* prefix). The second coda condition (9b) is similar to Steriade's (1992) claim that the release phase of plosives is lost: $A_0A_{f/max} \rightarrow A_0$ in coda position. This condition can again be satisfied in two ways: by deleting either ʔ or **h**. The choice appears to depend on a language specific parameter, but other factors may modify it. As the following show, Hungarian by default chooses to delete the **h** element of the coda. We will see below that there are cases when condition (9b) is satisfied by the deletion of ʔ or it is not satisfied at all.

The processes in (8) can be formalized as shown in (10):

(10) **Adaffrication processes**
(place and laryngeal specifications omitted, heads underlined)

a.	b.	c.	d.
t + s	t + t͡s	t͡s + s	t͡s + t͡s
x x	x x	x x	x x
| |	| |	| |	| |
h̲ h	h̲ h	h̲ h	h̲ h
| |	| |	| |	| |
ʔ≫	ʔ ʔ	ʔ≫	ʔ ʔ

The coda ʔ in (10a and c) spreads to the following onset position. The coda ʔ in (10b and d) is fine: it is supported by the following onset's ʔ element. (OCP probably merges the

[17] This is a simplification of the facts: the cases in (8a and b) occur practically obligatorily, but those in (8c) and (8d) are effected only with the increase of tempo. This will gain importance below.

[18] Coda here means the first position of the consonant cluster. I ignore the question whether this is in fact a coda position or some other type of lenition site.

two, but this has no bearing on the issue discussed.) The coda's **h** is deleted in all four cases. As a result the codas are all containing ʔ and the following onsets **ʔ.h**, that is [tts] (=[t͡sː]).

With the representations and conditions given, the gemination of [tt], (11a), can also be produced: the first [t] loses its **h**, becoming unreleased, the second is not changed, the result is a geminate [tː]. [s]-initial clusters, [st] and [st͡s], are not affected since their first element does not contain both **h** and ʔ (11b). Affricate+stop clusters do pose a problem, however. What we expect to happen here is that the **h** deletes and thus the first element becomes an unreleased stop: *t͡s+t→tː*. This is not what happens. If anything, it is the ʔ element that deletes (11c), but in careful speech nothing happens (11c′): *strucctoll* [-t͡st-]~[-st-] 'feather of an ostrich', *becstelen* [-ʧt-]~[-ʃt-] 'dishonest'. Comparing the representations in (11) with those in (10), there seems to emerge a condition, which is somewhat arbitrary in this framework,[19] stating that the **h** of simultaneous **h** and ʔ coda elements can only be deleted if it is not head or if the next segment's head is also **h**:

(11) **Further coda releases**

a.	t + t	b.	s +t/t͡s	c.	t͡s+ t	c′.	t͡s+ t
	x x		x x		x x		x x
	| |		| |		| |		| |
	~~h~~ h		h h		h h		h h
	| |		|		| |		| |
	ʔ ʔ		ʔ		~~ʔ~~ ʔ		ʔ ʔ

The Basque stop deletion and affricate spirantization rules seem to be governed by the same conditions as the Hungarian adaffrication rules. The condition in (9a) is not applicable in this case since there are no plosive+fricative clusters in the input (at least as far as Hualde's data are concerned). The constraint (9b) formulates is satisfied by the other option in Basque: whereas in Hungarian it was the **h** element that was deleted by default, in Basque it is ʔ.

(12) **Basque stop deletion and affricate spirantization**
(place and laryngeal specifications omitted, heads underlined)

a.	t + p	b.	t͡s+ t
	x x		x x
	| |		| |
	h h		h h
	| |		| |
	~~ʔ~~ ʔ		~~ʔ~~ ʔ

In the representation proposed the difference between the manner structure of a stop and an affricate is one of headedness. If an affricate loses its ʔ element, what remains is its head, **h**, interpreted as a fricative (12b). This is the case of /its-tegi/ → [isteɣi] 'dictionary'. When a stop loses the ʔ on the other hand, it loses its head,[20] and the remaining headless expression, h._, is uninterpretable in Basque, cf. /bat paratu/ → [baparatu] 'put one'. Following

[19] The state of affairs seems to call for an Optimality Theoretic account along the following lines: the coda conditions in (9) are violable constraints. A constraint DON'T-DELETE-HEAD is ranked higher, so when satisfying (9b) would entail deleting the head either nothing happens, thus fulfilling PARSE, or the ʔ element is deleted, violating the language specific parameter requiring that **h** be deleted of simultaneous coda ʔ and h. Somehow — the details are to be worked on — deletion of a head is more feasible if the following onset has the same head, cf. (10c and d). To further complicate the issue, it must be admitted that no change and deletion of ʔ is also possible in affricate+affricate clusters (e.g. *malaccomb* [-t͡sː-]~[-t͡st͡s-]~[-st͡s-] 'pig's ham'), implying a very intricate hierarchy of constraints.

[20] DON'T-DELETE-HEAD is probably lowly ranked in Basque.

a completely unrelated line of inquiry, Backley has proposed **h.**_ to be the representation of [h] (1993:315). This interpretation would explain both the Basque case, which lacks the segment [h], and the Yucatec Mayan examples, where homorganic stop+stop clusters become [h]+stop.

3.2 A problem with place

As we have seen the unary element proposal is far from being without its own disadvantages. A serious problem is caused by the fact that the adaffrication processes formalized in (10) hold only for coronals and palatals, not for labials and velars. There is no formal way of capturing these two places of articulation in this framework. Coronals may be assumed to contain an **R** element, while palatals have an **I**. Even if we further assume that palato-alveolars have both **R** and **I**, we can only involve the place elements if we posit a rule that makes reference to them: there is nothing in **R** (or **I**) in itself that should make coronals more prone to this process than fricatives and stops at other places of articulation. Though including place specification in the rule is the standard way of dealing with the phenomenon, rules involving but not affecting place elements are too strong a device and ought to be eliminated. The problem thus remains but it can be parallelled by a number of other facts that seem connected.

Greek stop+fricative (i.e., stop+[s]) clusters can only contain a noncoronal stop: [ps], [ks] vs. *[ts], and the same applies to morpheme-internal clusters of English and many other langauges. That coronal noncontinuants are less stable than noncoronals is also evidenced by nasal+continuant clusters. Whereas clusters like [ms] usually retain their nasal noncontinuant (or may even cause postnasal hardening in, say, Hungarian: *szomszéd* [som(p)seːd] 'neighbour'), in [ns], as we have seen above, the nasal dissolves into the preceding vowel. The instability of the coronal stop and affricate could well be the cause for their willingness to merge with the following strident fricative or affricate in adaffrication processes in Hungarian.

4 Conclusion

We have seen that phonological models of affricates have changed considerably during the past decades. The false predictions of contour segment analyses called for alternative representations. These likened affricates to stops either by saying that neither are contours (Schafer 1995) or that both stops and affricates are contour segments inasmuch as both have two root nodes (Steriade's works). This paper follows the first of these approaches since it does not distinguish stops and affricates on the one hand and other segments on the other in the number of root nodes they possess, but instead it is claimed that the two types of plosives differ in the dependency relations of their "stricture" elements: stops are noncontinuant (**ʔ**) headed, while affricates are headed by noise (**h**). Examining Hungarian adaffrication we have found that the development of well-formed codas seems to be governed by ranked constraints and that place specifications have their role in the behaviour of the stricture properties of these segments, the exact details remain a mystery.

References

Backley, Phillip. 1993. Coronal: the undesirable element. *UCL Working Papers in Linguistics* 5:301–323.

Brockhaus, Wiebke. 1995. *Final devoicing in the phonology of German.* Tübingen: Niemeyer.

Catford, John C. 1988. *A practical introduction to phonetics.* Oxford: Calendron.

Chomsky, Noam and Morris Halle. 1968. *The sound pattern of English.* New York: Harper & Row.

Halle, Morris and George N. Clements. 1983. *Problem book in phonology.* Cambridge, Mass.: MIT Press.

Harris, John. 1990. Segmental complexity and phonological government. *Phonology* 7:255-300.

Harris, John. 1994. *English sound structure.* Oxford: Blackwell.

Harris, John and Geoff Lindsey. 1995. The elements of phonological representation. In: Jacques Durand and Francis Katamba (eds.) *Frontiers of phonology: atoms, structures, derivations.* Harlow: Longman. 34–79.

Hoard, James E. 1971. The new phonological paradigm (review article on the SPE). *Glossa* 5 : 222–268.

Hualde, José Ignacio. 1988. Affricates are not contour segments. *Proceedings of WCCFL VII.* 143–157.

Jakobson, Roman, C. Gunnar M. Fant and Morris Halle. 1952. *Preliminaries to speech analysis. The distinctive features and their correlates.* Cambridge, Mass.: MIT Press.

Jones, Daniel. 1972. *An outline of English phonetics (9th ed.).* Cambridge: Cambridge University Press.

Kaye, Jonathan. 1990. 'Coda' licensing. *Phonology* 7 : 301–330.

Kaye, Jonathan. 1992. Do you believe in magic? The story of *s+C* sequences. *SOAS Working Papers in Linguistics and Phonetics* 2 : 293–313.

Kaye, Jonathan, Jean Lowenstamm and Jean-Roger Vergnaud. 1990. Constituent structure and government in phonology. *Phonology* 7 : 193–231.

Kenstowicz, Michael. 1994. *Phonology in generative grammar.* Cambridge, Mass.: Blackwell.

Laziczius Gyula. 1977. *Fonetika* [Phonetics (3rd ed.)]. Budapest: Tankönyvkiadó.

Lombardi, Linda. 1990. The nonlinear organization of the affricate. *Natural Language and Linguistic Theory* 8 : 375–426.

Maddieson, Ian. 1984. *Patterns of sounds.* Cambridge: Cambridge University Press.

Polgárdi Krisztina. 1991. Affricates as contour segments (in Hungarian and in English). MA dissertation, Eötvös Loránd University (ELTE), Budapest.

Rubach, Jerzy. 1994. Affricates as strident stops in Polish. *Linguistic Inquiry* 25 : 119–143.

Sagey, Elizabeth. 1986. The representation of features and relations in nonlinear phonology. PhD dissertation, MIT, Cambridge, Mass.

Schafer, Robin. 1995. Headedness in the representation of affricates. *The Linguistic Review* 12 : 61–87.

Siptár Péter. 1994. A mássalhangzók [The consonants]. In: Kiefer Ferenc (ed.) *Strukturális magyar nyelvtan 2. Fonológia* [A structural grammar of Hungarian 2. Phonology]. Budapest: Akadémiai Kiadó. 183–272.

Steriade, Donca. 1992. Segments, contours and clusters. Ms, UCLA.

Steriade, Donca. 1993. Closure, release, and nasal contours. In: Marie Huffman and R. Krakow (eds.) *Nasals, nasalization, and the velum. (Phonetics and phonology 5.)* San Diego, Ca.: Academic Press. 401–470.

Steriade, Donca. 1994. Complex onsets as single segments: the Mazateco pattern. In Jennifer Cole and Charles Kisseberth (eds.) *Perspectives in phonology.* Stanford, Ca.: CSLI. 203–291.

Szigetvári, Péter. 1996. Laryngeal contrasts and problematic representations. In: László Varga (ed.) *The Even Yearbook 2.* Budapest: Department of English Linguistics, Eötvös Loránd University (ELTE). 97–110.

Törkenczy, Miklós. 1989. Does the onset branch in Hungarian? *Acta Linguistica Hungarica* 39 : 273–292.

Theoretical Linguistics Program
Eötvös Loránd University (ELTE)
1014 Budapest, Színház utca 5–9, Hungary

szigetva@nytud.hu

Temporal dependencies and the interpretation of tenses*

Trón Viktor

Abstract

The present paper is a semantic study of some problems that arise in the treatment of natural language tenses. I will confine myself to the present and past tenses, especially focussing on their dependency in complex sentences in English and Hungarian. In order to give a formal account of the phenomena a semantic representation framework is developed that is an integration of two previously existing paradigms.

1 The relevant data

In this section I will briefly sketch some of the most important facts about complex tensed sentences. In particular, I will put the emphasis on recognising contrasts between different types of sentences and languages, such as the contrast between complex sentences with relative clauses and ones with complement clauses or the contrast between Hungarian and English examples of the latter construction. Also I shall discuss the possible interpretations of some English sentences that prove rather interesting.

1.1 Contrast between relative and complement clauses

Let us investigate the two sentences below. Both of the sentences contain two tensed clauses all of which are in the past tense. The difference between

*I am greatly indepted to Ágnes Bende-Farkas, who encouraged me and helped me in writing this article. I also benefited from discussions with László Kálmán.

them is that whereas the first one instantiate a complex sentence where the direct object is modified by a relative clause, in the case of the other example the verb takes the second clause as its complement.

(1) Victor knew the girl who was in the pub.
(2) Victor knew that the girl was in the pub.

We can clearly see that in (1) the interpretation of the two past tenses is independent, which is to say that they allow for the event times to be in any order with respect to each other. This, however, is not the case in (2). As the event time of the complement clause can not be wholly preceded by the event time of the main clause. It is apparent that the interpretation of the embedded clause past tense is somehow dependent on that of the matrix clause past tense. The above distinction then makes it clear that the interpretation of tenses is in this respect **structure dependent**. Any account of tenses should be armed to handle this phenomenon.

1.2 Contrast between Hungarian and English

The following two examples show that the behaviour of tenses diverge in languages. We see that even if under the same structural circumstances the interpretation of a past tensed clause embedded under a past tense clause is different in the case of Hungarian and English.

(3) Viktor tudta hogy Zsuzsanna terhes volt.
Viktor know [PAST, 3rd, SING] that Suzanne pregnant was
'Viktor knew that Suzanne had been pregnant.'
(4) Victor knew that Suzanne was pregnant.

Whereas the Hungarian sentence has the only reading when the time of *Suzanne's pregnancy* strictly precedes *Viktor's knowing it*, in English the two **event times** can also be **simultaneous**. This amounts to saying that the interpretation of tenses is not unified cross-linguistically. Other examples from Hungarian suggest that tenses are not necessarily evaluated against the **time of utterance** but against a **reference time** which might be other than the speech time in embedded sentences. Any account that can give a unified interpretation of tenses as well as can explain how the reference time is

bound by the matrix event time is ready to explain this behaviour of tenses in Hungarian and similar languages. Things, however, are far more complicated in the case of English as we can see in the next subsection.

1.3 Double-access constructions

In the next example we investigate a sentence where a present tense clause occurs as the complement of a past tense main verb.

(5) Viktor tudta hogy Zsuzsanna terhes.
Victor know PAST 3rd, SING that Suzanne pregnant
'Viktor knew that Suzanne was pregnant.'
(6) Victor knew that Suzanne is pregnant.

As we can already predict, the Hungarian sentence implies that the time of *Suzanne's pregnancy* is simultaneous with the time of *Victor's knowing it*,but not necessarily simultaneous with speech time. In other words the present tense in the embedded clause is interpreted relative to a reference time that is the event time of the main clause. In the English example we have an additional constraint on the interpretation of the sentence, namely the time of *Suzanne's pregnancy* must also **overlap** the speech time. Since on the one hand the embedded present must be simultaneous with the matrix event time and on the other hand it must be present relative to the time of utterance as well, this is traditionally referred to as a **double-access construction**. Constructions like this suggest the **deictic** nature of English embedded tenses, whereas the Hungarian examples manifest an **anaphoric** nature. There are problems, however, if we only assume that English tenses are always evaluated against the time of utterance, as we see from (2). Nevertheless I shall pursue this idea in the present paper and try to solve the problem just raised.

1.4 Further contrasts

So far we tried to state certain generalisations about the interpretation of tenses and suggested that in the case of certain constructions such as *past embedded in past* sentences we can have two readings that are different with respect to the order of the events denoted by the clauses. We might ask the question whether this is always so regardless of what kind of verb or verb

phrase occurs in a given sentence.

(7) Victor said that he ran three miles.
(8) Victor said he was running.

(1) shows that while the first can only be interpreted with a **past shifted** reading (2) can also have a **simultaneous** interpretation.

Also we can observe that double access constructions are *not* always available, as we see in the examples below.

(9) Sue thought that she is pregnant.
(10) *Sue thought that she walks a few miles.

This indicates that the interpretation of tenses is **meaning sensitive**[1] and thereby subject to complex semantic constraints which a successful theory of tenses should provide. Similarly, the exact characterisation of verb phrases that allow double access constructions should also be given.

2 Previous Accounts

Similar problems in connection with the temporal dependencies in complex sentences have already been adressed to in a lot of works in the literature. Among the papers that especially focussed on complex sentences are [Enç 87] and [Stowell 95b]. Enc's work developes a syntactic account for the contrast between relative and complement clauses based on a Government and Binding framework. The second contrast proves to be problematic for her, although she proposes an alternative analysis in terms of an altered definition of tense interpretation in languages like Hungarian or Russian. In [Stowell 95b] the first three contrast are explained in a syntactic theory of tenses developed in [Stowell 95a]. English tense morphemes are interpreted as *referential expressions* with the past tense morpheme exhibiting a past *polarity* property, similar to the well-known negative polarity. Although he can retain the universal interpretation of *tense predicates* but only at the cost of assuming that

[1]The term *meaning sensitive* may not be the best choice. It is rather the lexical properties of the verb as well as the aspectual properties of the embedded sentence that determine whether a double-access is available or not.

tense phrases have divergent characteristics cross-linguistically. Although the possibility of double-access constructions is accounted for by Stowell but the apparent contrast revealed in 1.4. remains unexplained in both approaches, which considerably reduces their explanatory force.

3 Event semantics and information states

In this section I will build up the apparatus needed for the semantic theory of tenses. In the first subsection I will present a brief overview of event semantics, one developed in [Krifka 89] and [Krifka 92], while in the second I will outline the theory of information states and verification of [Crouch 93]. In the third subsection I will try to integrate the two theories so that it serve to explain the tense phenomena that are of our concern.

3.1 Event Semantics

The Event semantics that I am going to use is basically the one developed in [Krifka 92]. He assumes that there are three distinct types of entities in the universe ; namely **objects**, **events** and **times**[2]. The structure of these is a **complete join semi-lattice** without a bottom element. So we have the operation $\sqcup$ on the set of all events as well as we have the following three relations: **part**, **proper part** and **overlap** ($\sqsubseteq$, $\sqsubset$ and $\circ$, respectively).

We make the additional assumption that the time lattice is atomic and we define the temporal order relation $\leq$ which is a linear order for time points. The relation between events and times are guaranteed by the function τ that maps from the extension of events to the extension of times. τ is a homomorphism with respect to the *join* operation.

For a syntactic framework [Krifka 92] uses a categorial-like representation with concatenation as the only syntactic operation on the one hand and function application as a semantic operation on the other. Verbs are taken as one-place predicates over the set of events. Here the syntactic arguments of verbs are treated with syntactic categorisation and they have no semantic counterparts. Merely for expository convenience I will diverge from Krifka

[2]In [Krifka 92] these are characterized by the predicates O, E and T, respectively. I also adopt the notation that these letters with indices are used for variables the value of which can only be a corresponding type of entity.

at this point and work with the simpler assumption that verbs are n-place predicates over the set of events and objects. In other words the syntactic arguments of the verb are taken as its semantic arguments with an additional event argument. In the case of a derivation, after all free variables are bound we get a property of events, this Krifka calls the **sentence radical**. It is then a **sentence mood operator** that turns this into a sentence by virtue of binding the event variable with an existential quantifier. The sentence in this form can be evaluated according to traditional model-theoretic truth definitions.

We have to discuss the account of tenses at this point. I will treat tenses as sentential operators that predicate over the **temporal trace** of the events. Along these lines we can define the syntax and semantics of tense morphemes:

Definition 1

1. PAST S/S $\lambda P \lambda e.P(e) \wedge Past(e)$
2. PRESENT S/S $\lambda P \lambda e.P(e) \wedge Pres(e)$

Where Past and Present are simply ordering predicates over times:

$$[\![Past(e)]\!]^{t_0} = \begin{cases} 1 & \text{if } \tau([\![e]\!]) \prec t_0 \\ 0 & \text{otherwise} \end{cases}$$

$$[\![Pres(e)]\!]^{t_0} = \begin{cases} 1 & \text{if } \tau([\![e]\!]) \circ t_0 \\ 0 & \text{otherwise} \end{cases}$$

where $\prec$ is the *strict precedence* relation based on the linear order $\leq$.

3.2 Information states and verification time

In [Crouch 93] a semantic theory is outlined to handle the relations between the members of conditional sentences. In this section I will briefly summerise the basic ideas. Recognising the obvious epistemic asymmetry between present and past tense on the one hand and future tenses on the other, [Crouch 93] splits update into **assertion** and **verification**, but in the case of simple sentences the time of assertion and verification coincide, as simple declarative sentences are used to convey verified information.

Information states are taken as primitives and his information model is a tuple.

(11) Information model

$M = \langle S, E, I, \sqsubseteq_t, \leq, \trianglelefteq, V \rangle$
where S is a set of information states
E is a set of time points and periods
I is a subset of E containing only time points
$\sqsubseteq_t$ is a relation of information extension in $S \times I \times S$
(transitive and reflexive for any $t \in I$)
$<$ is a linear order of temporal precedence on E
$\trianglelefteq$ is a relation of temporal inclusion on E
V is a valuation function for atomic sentence letters

[Crouch 93] come up with a set of well motivated constraints that the information model just defined has to satisfy[3]. I adopted his notational convention that v stands for verification time, p for an atomic sentence letter and t_p for event time[4].

1. Monotonicity of verification:
   If $V(s, v, p, t_p) = 1$ then for all time indices $v' > v$, $V(s, v', p, t_p) = 1$

2. Monotonicity of information extension:
   If $s_1 \sqsubseteq s_2$, then

   (a) if $V(s_1, v, p, t_p) = 1$, then $V(s_2, v, p, t_p) = 1$

   (b) if $V(s_1, v, p, t_p) = 1$ for some point v, then $V(s_2, v', p, t_p) = 1$ for some time point v'

3. (No) Fore-verification
   If $V(s, v, p, t_p) = 1$, then $v \geq t_p$

The truth conditions of tense operators are given with the help of the valuation function and evaluated with respect to an additional temporal parameter a

[3]I left out three of these, namely *Richness*, *Convergence of Verification* and *No Absurdity* for the sake of simplicity and because they are irrelevant in the present approach.

[4]In [Crouch 93] the language contains propositions that are combined with operators and the usual logical connectives. Here the term *event* is used in a different sense from Krifka and t_p stands for the time interval associated with p.

which stands for the assertion time.

(12) Operator definitions

$s, a, v, e \models p$ iff $V(s, v, p, e) = 1$
$s, a, v, e \models past(\phi)$ iff there is some $e' < a$ such that $s, a, v, e' \models \phi$
$s, a, v, e \models pres(\phi)$ iff there is some $e' \geq a$ such that $s, a, v, e' \models \phi$

3.3 Integrating the two paradigms

If we want to integrate event semantics with information semantics there is a wish to reduce the number of primitives in the system. In his section I try to construct an information model out of events. If we adopt the event and time structure discussed in section 3.1 we can think of information states as subsets of the set of all events with a temporal parameter specified. In [Crouch 93] the truth of a given formula is evaluated against four parameters one of which is the time of verification. If we want to incorporate verification information into our definition of information states we additionally have to specify a set of events which is a subset of the events asserted in an information state and contain the ones that have already been verified. This is characterized by the predicate Ver over events, the extension of which is always the current verification set of the information state. Based on what has been sad so far we can define our information states.

Definition 2 $\mathcal{S} \subseteq \mathcal{P}(\mathcal{E}) \times \mathcal{T} \times \mathcal{P}(\mathcal{E})$
where $\mathcal{S}$ is the set of all information states, $\mathcal{E}$ is the extension of all events, $\mathcal{T}$ is the set of times in our model.

Obviously we do not want the two sets to coincide in the general case, we have to impose constraints on the set of possible information states. We saw in [Crouch 93] that he temporal parameter of an information state can only be a timepoint not a period.

(13) if $s = \langle E, t, V \rangle \in \mathcal{S}$, then $t \in ATOM(\mathcal{T})$

Here $ATOM(\mathcal{T})$ means the subset of $\mathcal{T}$ containing only the atomic times (T_a in [Krifka 92]).

The definition of information extension seems quite natural here given that we can base it on set-theoretic *subset* relation.

Definition 3 $s = \langle E, t, V\rangle \sqsubseteq s' = \langle E', t', V'\rangle$ iff $E \subseteq E'$ and $t \leq t'$ and $V \subseteq V'$

It is now interesting to see whether we can adapt [Crouch 93]'s constraints to our information model. There it is the valuation function that says whether a given *proposition* is true with respect to an information state, a verification time and an event time. In our case it is the interpretation of $Ver(e)$ that gives us the same information with the exception that we do not have propositions only events. Now we can look at the reformulated constraints and see if they hold.

1. Monotonicity of verification:
   If $[\![Ver(e)]\!]^{\langle E,t_1,V\rangle} = 1$, then for all time indices $t_2 > t_1$,
   $[\![Ver(e)]\!]^{\langle E,t_2,V\rangle} = 1$

2. Monotonicity of information extension:
   If $s_1 \sqsubseteq s_2$, then $[\![Ver(e)]\!]^{s_1} = 1$, then $[\![Ver(e)]\!]^{s_2} = 1$

One can easily see that these constraints automatically hold because of the monotonicity of the *element of* relation with respect to the *part of* relation. The remaining two constraints are assured in the following definitions.

Definition 4 (No) Fore-verification
$s = \langle E, t, V\rangle$, if $e \in V$, then $\tau(e) \preceq t$ where $\preceq$ is the *precedence* relation based on the linear order $\leq$.

Additionally we have to stipulate that information states are closed under the part of relation on events, this is supposed to capture the intuition that if an information state contains the event of *John's watching the TV from 8:30 to 10:30* as a verified event then it contains its subevents, e.g. *John's watching the TV between 9 and 10* also as a verified event. This is formalised below:

Definition 5 $s = \langle E, t, V\rangle$, if $e \in E$, then for any $e' \sqsubseteq e$, $e' \in E$ and similarly if $e \in V$, then for any $e' \sqsubseteq e$, $e' \in V$

So far we have not talked about natural language sentences. How does our information states tell us anything about the truth of a sentence? Sentences are taken as existential statements about events; this is achieved via an operator that binds the free variable in the sentence radical. If we reinterpret the sentence mood operator for normal declaratives, we are supposed to represent that the information conveyed by the sentence is already verified. For this we

need to split our sentence mood operator into a conjunction which tells us that there is an event denoted by the sentence radical which is verified with respect to the information state given[5].

Definition 6 OP S'/S $\lambda P.\exists e(P(e) \wedge Ver(e))$

As we are concerned here with complex sentences in which there are clauses in the embedded argument position of certain predicates we now turn to the question of their representation. Syntactically these predicates are taken as ones taking an S type argument, which is the sentence radical; semantically speaking, however, the evaluation of this sentence depends on the matrix event denoted by the actual main predicate or at least on its temporal trace. On the other hand, we saw that in English sentences the complement clause can contain the speaker's additional information -temporal in character - that is not to be present in the semantic argument of the predicate.[6]. We definitely have to make an attempt at representing this double dependency by formal means. We can see in Hungarian examples that such additional information on the part of the speaker can never be made use of. In other words, the dependency on the main predicate is *syntactically preencoded* while in English it is not. This contrast is interpreted as a difference in the semantics of predicates taking a clausal complement[7].

Definition 7 If P is a predicate taking a clausal complement, then

1. In English:
$$[\![P(e,s)]\!]^s = \begin{cases} 1 & \text{if } \bigvee e' \text{such that} e' \in [\![s]\!] \text{ and} \\ & \langle [\![e]\!]^s, e' \rangle \in [\![P]\!]^s \\ 0 & \text{otherwise} \end{cases}$$

2. In Hungarian:
$$[\![P(e,s)]\!]^s = \begin{cases} 1 & \text{if } \bigvee e' \text{such that} e' \in [\![s]\!]^{\langle E,\tau([\![e]\!]^s),V\rangle} \text{ and} \\ & \langle [\![e]\!]^s, e' \rangle \in [\![P]\!]^s \\ 0 & \text{otherwise} \end{cases}$$

[5]This formalism in principle enables us to represent further types of sentences, i.e, where the 'sentence' is not meant to convey verified information, but only assertions and maybe also modal statements.

[6]This is also the case in sentences like: *Peter said he will come.*

[7]It should be noted that in this interpretation *intensionality* is treated in a rather rough manner. This is because it is irrelevant for the present purposes; things, however, can be altered if we introduce a *possible worlds* framework.

3.4 Verifiability

It is interesting to observe that almost all of the predicates taking a clausal complement are related to information states. The verbs *believe*, *claim*, *know*, *say* or *think* all contain an implicit reference to an information state somehow related to their subjects. It is therefore not surprising that the predicates above impose certain constraints of the interpretation of their complements. This is exclusively the result of the constraints on our information models that also apply to the information states implicitly referred to. As this is not present in the semantic representation of the sentence, we have to state this in a form of a postulate.

(14) **The sentential complement postulate**: The event denoted by the predicate in the complement clause has to have a faithful subevent that is verifiable at the time of the matrix event.

I call a subevent e' of an event e **faithful** with respect to a predicate P ($e \sqsubseteq_P e'$) if except for its temporal trace P can be predicated of both e and e'. Here **verifiability** is to be understood very strictly, in accordance with the *No fore-verification* constraint an event e is verifiable at time t only if $e \preceq t$.

The above postulate amounts to saying that in the sentence *John thought that he was running* the event of *John's running* has to have a faithful subevent verifiable at the time of *John's believing it*[8]. If we wish to build up a more sophisticated semantic theory we might just as well try to represent this implicit knowledge in the actual semantic representation of the verb's meaning. With this in mind we can redefine the interpretation of predicates taking a sentential complement.

Definition 8 If P is a predicate taking a clausal complement, then

1. In English:

$$[\![P(e, \lambda x_e P'(x_e))]\!]^s = \begin{cases} 1 & \text{if } \bigvee e', e'' \text{ such that} \\ & e' \in [\![s]\!]^s \text{ and } \langle [\![e]\!]^s, e' \rangle \in [\![P]\!]^s \text{ and} \\ & e'' \sqsubseteq_s e' \text{ and } \tau(e'') \preceq \tau([\![e]\!]^s) \\ 0 & \text{otherwise} \end{cases}$$

[8]Note that his is only true to sentences that do not imply a modal context. In the case of e.g. *John thought that he should become a doctor.* things are more complicated and not dealt with here.

2. In Hungarian:

$$[\![P(e, \lambda x_e P'(x_e)]\!]^s = \begin{cases} 1 & \text{if } \bigvee e', e'' \text{ such that} \\ & e' \in [\![s]\!]^{\langle E, \tau([\![e]\!]^s), V\rangle} \text{ and} \\ & \langle [\![e]\!]^s, e'\rangle \in [\![P]\!]^s \text{ and} \\ & e'' \sqsubseteq_s e' \text{ and } \tau(e'') \preceq \tau([\![e]\!]^s) \\ 0 & \text{otherwise} \end{cases}$$

4 Explaining the contrasts

In this section I attempt to use the semantic framework developed above to account for the contrasts discussed in the first section. If we look at our rules that determine the temporal dependencies in sentences we can see that tenses have universal interpretation, and in the case of matrix tenses their evaluation only depends on the information state (in fact, only its temporal parameter). When a relative clause occurs in a subordinate position modifying a constituent we would predict that its tense is evaluated relative to the utterance time similarly to the matrix tense. In the case of complement clauses, however, their being arguments of the verb causes them to be dependent in some ways on the predicate. We saw that there are two ways to this dependency: one is purely syntactic, when the interpretation of tenses is relativized to the matrix event time. In Hungarian it plays no role in the interpretation of the tenses whether we use a specific type of predicate in the embedded clause or another; this also confirm the merely syntactic nature of this phenomena. The other sort of dependency is only due to our knowledge of what the actual matrix predicates mean: the fact that they are closely related to information states. In fact, this rule only constrains the set of possible readings of sentences and they are sensitive to certain properties of the embedded clause (the predicate it contains).

The usual format of a sentence containing an embedded clause in complement (e.g., (15)) position looks like (16).

(15) Victor thought that Sue was in the pub.
(16) $\exists e(tense(e) \wedge P(e, \lambda e'.P'(e')) \wedge Ver(e))$

Uttered in an information state $s = \langle E, t, V\rangle$ - substituting past tense - it

is interpreted in Hungarian as (17)[9].

(17) $\bigvee e, e', e'' \in \mathcal{E}$ $\tau(e) \preceq t$ and $\tau(e') \prec \tau(e)$ and $\langle e, e' \rangle \in [\![P]\!]^s$ and $e' \in [\![P']\!]^s$ and $e'' \in [\![P']\!]^{\langle E, \tau(e), V \rangle}$ and $\tau(e'') \preceq \tau(e)$ and $e \in V$

We see in (17) that the condition $\tau(e'') \preceq \tau(e)$ does not restrict the truth conditions of the sentence since it already follows from $\tau(e') \prec \tau(e)$[10].

Things are far more complicated in the case of English, where it is not a purely syntactic dependency that proves to be relevant. In what follows I will focus on the analysis of this latter type of dependency and the ways how it determines the readings of English sentences.

(18) $\bigvee e, e', e'' \in \mathcal{E}$ $\tau(e) \preceq t$ and $\tau(e') \preceq t$ and $\langle e, e' \rangle \in [\![P]\!]^s$ and $e' \in [\![P']\!]^s$ and $e'' \in [\![P']\!]^s$ and $\tau(e'') \preceq \tau(e)$ and $e \in V$

In (18) we can see that the existence of a suitable subevent is not immediately obvious as was the case in the Hungarian sentence. It is interesting to see how the evaluation depends on the actual choice of P' the embedded predicate in the sentences.

Predicates over events can be classified according to whether they are closed under the *part* relation on events. This distinction is made explicit in [Krifka 92], where predicates which manifest the closedness property are said to be **divisive**. Those predicates in the extension of which there are no two events that stand in the *proper part* relation are called **quantized**[11].

Definition 9 $P \in$ **DIV** iff for every pair of events e ande' if $P(e)$ and $e' \sqsubset e$, then $P(e)$

Definition 10 $P \in$ **QUA** iff for every pair of events e and e' if $P(e)$ and $e' \sqsubset e$, then $\neg P(e')$

[9] I left out the details of the exact syntactic derivation as well as the detailed calculation of the truth conditions represented in (17).

[10] It should be added that in the case of present tense embedded in past the existence of a subevent might be problematic. If it turned out that the predictions on Hungarian proved to be correct, than it would confirm the universality of (14), but for the time being I remain silent about this issue.

[11] Here P obviously stands for all entities of type e/S and not necessarily for atomic predicates. In the present paper I am not concerned with the issue how these properties can be derived (probably compositionally) from the lexical properties of the actual predicate and its arguments.

Generally speaking, predicates denoting **states**[12], such as *be pregnant*, *be running* or *love someone* are examples of divisivity; on the other hand, predicates denoting complex events, such as *reach the top*, *run a mile* or *write a letter* are quantized. It is easy to see that these are extremes with respect to the *faithful subevent* constraint. If we have a divisive predicate in the complement clause then every subevent of any event in its extension is faithful. On the other hand in the case of a quantized predicate the events in its extension have no proper parts that are faithful. This adds up to saying that the truth value of (18) strongly depends on whether the embedded predicate is divisive or quantized.

From (18) it is clear that a necessary condition for the sentence to be true is to find a subevent of e' that is faithful and that precedes the time of the matrix event. In the case of quantized predicates occurring in the complement clause the only candidate is the whole e'' itself as an extreme case of the *part* relation as we saw above. Now if there is a present tense in the complement clause and a past tense in the main clause we get (19)which is apparently unsatisfyable.

(19) $\bigvee e, e' \ldots \tau(e) \prec t \ldots\ldots \tau(e') \circ t \ldots\ldots \tau(e') \preceq \tau(e) \ldots$

This predicts that sentences like (20)are semantically anomalous, uninterpretable.

(20) *John told me yesterday that he builds a house.
(21) John told me yesterday that he ate three apples.

In (21) there is a past tense complement clause with a quantized predicate. The sentence can only be true if the embedded event e' precedes e the matrix event - though not necessarily strictly.

(22) $\bigvee e, e' \ldots \tau(e) \prec t \ldots\ldots \tau(e') \prec t \ldots\ldots \tau(e') \preceq \tau(e) \ldots$

(22) correctly predicts that (21) can never be interpreted with the event of *John's eating three apples* not being completed by the time of *telling it.*

[12]Here states are to be understood in the broad sense, i.e., in the sense of [Moens 87] where it also includes progressive states, habitual states.

Now we can examine how divisive predicates behave in complement clauses. As we saw that in this case every subevent is faithful, the truth of the sentence only depend on the position of $\tau e'$. First in (23)there is a present tense in the embedded clause that contains the divisive predicate.

(23) Victor knew that Sue is pregnant.
(24) $\bigvee e, e', e'' \ldots \tau(e) \prec t \ldots\ldots \tau(e') \circ t \ldots\ldots e'' \sqsubseteq e'$ and $\tau(e'') \preceq \tau(e) \ldots$
(25) $e' \circ t$ and $e' \circ \tau(e)$

It is obvious in (24) that the sentence can only be true if there is a subevent e'' of the event e' such that $e'' \preceq \tau(e)$; this condition is equivalent to (25), which amounts to saying that the event denoted by the complement clause has to overlap t (now) as well as the time of the main clause event. This is exactly the condition of a *double-access* construction that we informally discussed in the first section. Now, we can predict that these are available with divisive predicates and not with quantized predicates in the complement clause.

(26) Sue knew that she was eating an apple.
(27) $\bigvee e, e', e''(\tau(e) \prec t \ldots\ldots \tau(e') \prec t \ldots\ldots e'' \sqsubseteq e'$ and $\tau(e'') \preceq \tau(e))$

In (27) we can, however, see how a simultaneous interpretation is possible in sentences like (26). The complement event e' has to be in the past but only some of its subevents has to precede the main clause event time $\tau(e)$.

5 Conclusions

In the present paper a semantic representation language is developed which integrates some aspects of update semantics with that of event semantics. Within this framework I discussed some problems related to natural language temporal phenomena, in particular the temporal dependencies of complex sentences with present and past tense. The relevant contrasts raised in the introductory part were all captured and probably accounted for, whereby I achieved a considerable coverage of the issue. Further applications and connections with other phenomena, such as aspectuality and modality, are still subject to research, which could help the theory to gain a better implement-

ation as well as dispose of its somewhat stipulative flavour.

References

[Krifka 89] Krifka, M., *Nominal Reference, Temporal Constitution and Quantification in event semantics.* In R.Bartch et al. (eds.) Semantics and Contextual expression, Foris, Dordrecht, 1989.

[Krifka 92] Krifka, M., *Thematic relations.* In Sag, I./Szabolcsi, A. (eds.) Lexical Matters. CSLI lecture notes, Stanford, 1992.

[Crouch 93] Crouch, R.,

[Moens 87] Moens, M., *Tense, Aspect and Temporal Reference*, Edinburgh, 1987, (pp. 38-69).

[Stowell 95a] Stowell, T., *The Phrase Structure of Tense*, In L.Zaring and J. Rooryck (eds.), Phrase Structure and The Lexicon, Kluwer.

[Stowell 95b] Stowell, T., *What do the Present and Past Tenses Mean?*, In P.Bertinetto et al (eds.), Temporal Reference, Aspect and Actionality', Vol. 1.:'Semantic and Syntactic Perspectives.', Rosenberg and Sellier, Torino, 1995.

[Enç 87] Enç, M., *Anchoring Conditions for Tense*, Linguistic Inquiry, Vol.18, Number 4., 1987, (pp. 633-657)

Double Negation in Dynamic Predicate Logic*

Károly Varasdi

Abstract

In this paper I am going to consider questions concerning double negation in Dynamic Predicate Logic and show that the lack of a dynamic version of negation is a necessary feature of the system.

1 The Problem

Consider the following three mini-discourses:

(A) *John owns a car. It is red, and it is parked in front of his house.*

(B) **John doesn't own a car. It is red, and it is parked in front of his house.*

(C) *It is not true that John doesn't own a car. It is red, and it is parked in front of his house.*

There is empirical evidence that while (B) is unacceptable, both (A) and (C) are accepted as sensible by most speakers. This phenomenon should be tackled by theories of discourse. In what follows I am going to consider if one such theory, Groenendijk and Stokhof's Dynamic Predicate Logic (DPL)[1], can cope with the above problem, and prove that DPL is not suitable for successfully dealing with it. (Groenendijk and Stokhof admitted that the version they proposed did fail to cope with the above problem[1] but expressed their hope that later work would overcome this imperfection.)

*I am greatly indebted to László Kálmán for his help in preparing the first version of this text. I am also grateful to Ágnes Bende-Farkas who taught me how to use LaTeX.

[1]Discourse (C) is their example; see Section 5.1 in [1].

2 The Basic Ideas of DPL

DPL was proposed by Groenendijk and Stokhof as an instrument which can tackle donkey sentences and anaphoric relations compositionally. Earlier proposals, such as Hans Kamp's Discourse Representation Theory, could handle these issues at the expense of giving up compositionality.

The basic idea of DPL is to treat meaning "dynamically". While the standard approach to sentence meaning is to equate it with truth conditions, the dynamic approach looks upon meaning as the ability to change the information state of the hearer. The utterance of a sentence brings the hearer from a certain information state to another one. Using computer terminology, one can say that a sentence is a program which, when executed, brings the hearer to another information state. In DPL, information states are identified with assignments of objects to variables, which is sufficient for coping with the problem of cross-sentential anaphoras. Whereas the interpretation of a formula in the standard semantics of predicate logic is the set of those assignments that verify the formula, in DPL the interpretation of a formula consists of ordered pairs of assignments: A pair $\langle g, h \rangle$ is in the interpretation of a formula ϕ iff when ϕ is evaluated ("executed") with respect to g, h is a possible outcome of the evaluation procedure.

3 Formal Details

3.1 Syntax

The syntax of DPL is that of ordinary First Order Predicate Logic.

$$
\begin{aligned}
&L = \langle LC, Con, Var, Term, Form \rangle, \\
&LC = \{(,), \neg, \wedge, \vee, \rightarrow, \exists, \forall, =\}, \\
&Con = Con_{ind} \cup Con_{pred}, \\
&Con_{ind} = \bigcup_{i \in \omega} \{a_i\}, \\
&Con_{pred} = \bigcup_{k \in \omega} Con^{(k)}_{pred}, \\
&Con^{(k)}_{pred} = \{P^{(k)}_i\}_{i \in \omega}, \\
&Var = \{x_i\}_{i \in \omega},
\end{aligned}
$$

$$
\begin{array}{lll}
Term = Var \cup Con_{ind}, & & \\
Form: & (i) & t_0, \cdots, t_{k-1} \in Term, P^{(k)} \in Con_{pred} \Rightarrow \\
 & & \Rightarrow P^{(k)}(t_0, \cdots, t_{k-1}) \in Form, \\
 & (ii) & t_1, t_2 \in Term \Rightarrow (t_1 = t_2) \in Form, \\
 & (iii) & \phi \in Form \Rightarrow \neg\phi \in Form, \\
 & (iv) & \phi, \psi \in Form \Rightarrow \phi \wedge \psi \in Form, \\
 & (v) & \phi, \psi \in Form \Rightarrow \phi \vee \psi \in Form, \\
 & (vi) & \phi, \psi \in Form \Rightarrow \phi \rightarrow \psi \in Form, \\
 & (vii) & \phi \in Form, x \in Var \Rightarrow \exists x\phi \in Form, \\
 & (viii) & \phi \in Form, x \in Var \Rightarrow \forall x\phi \in Form.
\end{array}
$$

3.2 Semantics

A model $\mathcal{M}$ is an ordered pair $\langle D, F \rangle$, where D is a non-empty set of individuals, F an interpretation function, having as its domain the individual constants and predicates in the following manner: If $a \in Con_{ind}$, then $F(a) \in D$; if $P^{(k)} \in Con^{(k)}_{pred}$, then $F(P^{(k)}) \subseteq D^k$.

An assignment g is a function assigning an individual to each variable:

$$g \in {}^{Var}D,$$

while $g[x]h$ is defined this way:

$$g[x]h \Longleftrightarrow \bigwedge_v (v \in Var \setminus \{x\} \Rightarrow g(v) = h(v)).$$

Finally, if G denotes the set of all assignments, then $[\![\cdot]\!]^{\mathcal{M}}_g \subseteq G \times G$ is the function assigning semantic values to the formulas in the following way[2].

$$
\begin{array}{l}
t \in Var \Rightarrow [\![t]\!]_g = g(t), \\
t \in Con_{ind} \Rightarrow [\![t]\!]_g = F(t),
\end{array}
$$

[2] I will drop subscripts and superscripts wherever this can be done without misunderstanding.

(i') $P^{(k)} \in Con^{(k)}_{pred}, t_1, \cdots, t_n \in Term \Rightarrow [\![P^{(k)}(t_1, \cdots, t_n)]\!] =$
$= \{ \langle g, h \rangle \mid h = g \,\&\, \langle [\![t_1]\!]_h, \cdots, [\![t_n]\!]_h \rangle \in F(P^{(k)}) \},$

(ii') $t_1, t_2 \in Term \Rightarrow [\![t_1 = t_2]\!] = \{ \langle g, h \rangle \mid h = g \,\&\, [\![t_1]\!]_h = [\![t_2]\!]_h \},$

(iii') $\phi \in Form \Rightarrow [\![\neg\phi]\!] = \{ \langle g, h \rangle \mid h = g \,\&\, \sim \bigvee_k : \langle h, k \rangle \in [\![\phi]\!] \},$

(iv') $\phi, \psi \in Form \Rightarrow [\![\phi \wedge \psi]\!] = \{ \langle g, h \rangle \mid \bigvee_k : \langle g, k \rangle \in [\![\phi]\!] \,\&\, \langle k, h \rangle \in [\![\psi]\!] \},$

(v') $\phi, \psi \in Form \Rightarrow [\![\phi \vee \psi]\!] = \{ \langle g, h \rangle \mid h = g \,\&\, \bigvee_k : \langle h, k \rangle \in [\![\phi]\!] \vee \langle h, k \rangle \in [\![\psi]\!] \},$

(vi') $\phi, \psi \in Form \Rightarrow$
$\Rightarrow [\![\phi \rightarrow \psi]\!] = \{ \langle g, h \rangle \mid h = g \,\&\, \bigwedge_k (\langle h, k \rangle \in [\![\phi]\!] \Rightarrow \bigvee_j : \langle k, j \rangle \in [\![\psi]\!]) \},$

(vii') $\phi \in Form \Rightarrow [\![\exists x \phi]\!] = \{ \langle g, h \rangle \mid \bigvee_k : k[x]g \,\&\, \langle k, h \rangle \in [\![\phi]\!] \},$

$(viii')$ $\phi \in Form \Rightarrow [\![\forall x \phi]\!] = \{ \langle g, h \rangle \mid h = g \,\&\, \bigwedge_k (k[x]h \Rightarrow \bigvee_m : \langle k, m \rangle \in [\![\phi]\!]) \}.$

Not all logical constants in DPL are going to be discussed, only those relevant in our discussion. These are *conjunction* and *negation*.Their definitions are reproduced below.

Definition 1 $[\![\phi \wedge \psi]\!] = \{ \langle g, h \rangle \mid \bigvee_k : \langle g, k \rangle \in [\![\phi]\!] \,\&\, \langle k, h \rangle \in [\![\psi]\!] \}$

Definition 2 $[\![\neg\phi]\!] = \{ \langle g, h \rangle \mid h = g \,\&\, \sim \bigvee_k : \langle h, k \rangle \in [\![\phi]\!] \}$

Let me now reinterpret what has been said in terms of concepts from the field of binary relations so that it will be easier to see the characteristics of the system.

The interpretation of a formula ϕ is a binary relation on G:

(1) $$[\![\phi]\!] \subseteq G \times G.$$

Let Φ be the set of all interpretations:

(2) $$\Phi = \mathcal{P}(G \times G),$$

while E is the identity relation on G:

(3) $$E = \{ \langle g, g \rangle \mid g \in G \}.$$

Now we can make an observation about the conjunction as construed in DPL which will be of importance later.

The definition of the conjunction can be reinterpreted as the operation of *composition* between relations:

(4) $$[\![\phi \wedge \psi]\!] = [\![\phi]\!] \circ [\![\psi]\!].$$

As for negation, its definition makes it clear that for any formula ϕ, the following holds:

(5) $$[\![\neg\phi]\!] \subseteq E.$$

This means that E is closed with respect to negation. As a consequence, double negation does not in general return with the semantic value of the original formula, except when it is already part of E. The interpretation of an existentially quantified formula is not of this kind. In other words, the system does predict a difference between an existentially quantified formula and its double negated form, which is a questionable feature[3].

4 The Problem Generalized

However, is it possible to define negation in such a way that double negation is always cancellable? Within the frames of DPL the answer is "no" for the following reasons.

Besides the cancellability of double negation, negation should satisfy the following condition:

(6) $$\bigwedge_{\phi}([\![\phi \wedge (\neg\phi)]\!] = \emptyset),$$

that is, a contradiction has to have the empty set as its interpretation.

The interpretation of the logical operator N we seek is a function from Φ, the set of all interpretations, onto Φ itself. N should satisfy the following condition expressing cancellability:

(7) $$\bigwedge_{\phi}([\![N(N(\phi))]\!] = [\![\phi]\!]),$$

which implies that the interpretation of N (denote it with $\mathbf{N}$) should be a bijective function. Condition (6) can be rewritten on the basis of (4) as

[3] "Next, example [C] shows that we need a dynamic version of negation as well, for which the law of double negation holds again. In fact, it can be expected that a suitable dynamic version of negation is all we need" (See Section 5.1 in [1]).

follows:

(8) $$\bigwedge_{[\![\phi]\!]} ([\![\phi]\!] \circ \mathbf{N}([\![\phi]\!]) = \emptyset).$$

However, this turns out to be inconsistent with the bijectivity of **N**. For the latter means that

(9) $$\bigwedge_{[\![\phi]\!]} \bigwedge_{[\![\psi]\!]} ([\![\phi]\!] \neq [\![\psi]\!] \Rightarrow \mathbf{N}([\![\phi]\!]) \neq \mathbf{N}([\![\psi]\!])),$$

which does not hold in general. The reason for this lies in the following property of composition.

(10) $$\bigwedge_{[\![\phi]\!]} ([\![\phi]\!] \circ \mathbf{N}([\![\phi]\!]) = \emptyset \iff \bigwedge_{h} (h \in Ran([\![\phi]\!]) \Rightarrow h \notin Dom(\mathbf{N}([\![\phi]\!]))))$$

Proof.
The "only if" direction is obvious by the definition of composition. The "if" direction is also easy. If $[\![\phi]\!] = \emptyset$, then we have nothing to do. So let us suppose that $[\![\phi]\!] \neq \emptyset$; then there is at least one $\langle g, h\rangle \in G \times G$ (where G is the set of all assignments) such that $\langle g, h\rangle \in [\![\phi]\!]$, i.e. $h \in Ran([\![\phi]\!])$. Now if for some k in G $\langle h, k\rangle \in \mathbf{N}([\![\phi]\!])$, then $\langle g, k\rangle \in ([\![\phi]\!] \circ \mathbf{N}([\![\phi]\!]))$ also holds. By contraposition we get that if $[\![\phi]\!] \circ \mathbf{N}([\![\phi]\!]) = \emptyset$, then there is no such k in G that $\langle h, k\rangle \in \mathbf{N}([\![\phi]\!])$, which in turn means that h is not in the domain of $\mathbf{N}([\![\phi]\!])$. $\mathcal{Q}.\mathcal{E}.\mathcal{D}$.

According to (10), for those $[\![\phi]\!]$s with $Ran([\![\phi]\!]) = G$ it is only the empty relation with which the right-product will be the empty set (as is required by (8)). But it is clear that it is in general possible to find a pair $[\![\phi]\!] \neq [\![\psi]\!]$ so that $Ran([\![\phi]\!]) = Ran([\![\psi]\!]) = G$. However, **N** cannot match different elements to these, so it cannot satisfy bijectivity. This, in turn, proves that it is not possible to define an operator N such a way that it satisfies the conditions demanded by negation.

5 Conclusion

The above considerations show that within the frames of DPL it is not possible to define negation in a way that universally satisfies the law of double negation. Consequently, the system cannot possibly tackle for theoretical reasons discourses like (C) under any improvement.

References

[1] Groenendijk, J. & Stokhof, M., "Dynamic Predicate Logic", 1991; In: Linguistics and Philosophy, Vol. 14., pp. 39 – 100.