Definiteness as a feature relevant to the conjugation of verbs in the Hungarian language

Neil Epstein
Swarthmore College
1998
Acknowledgements

Writing this thesis was not an entirely solitary event. I received invaluable encouragement and suggestions from Patrick Runkle, Kristen Vitalich, David Andrade, Thomas Kornack, Jennifer Freeman, Tisha Rajendra, Katherine Entigar, Elaine Huang, Na'im Tyson, Sandy Lin, Sarah Piatt (the other Swarthmore and Bryn Mawr senior linguistics majors of 1997-98), and George Gibbard (a first-year student), as well as from Professors Donna Jo Napoli, Tom Purnell, and Ted Fernald. I'm also very grateful to Professor Purnell for bringing to my attention the two articles of Farkas' that I used, and to both him and Professor Napoli for reading the thesis, and for all their organizational, topical, and theoretical suggestions. It was Professor Napoli, for instance, who suggested the idea of a tree to represent how OC gets assigned. If I forgot anybody, I apologize; it doesn't mean you're not appreciated!
Introduction

The purpose of this article is to say as much as possible about the Hungarian "Objective Conjugation". Hungarian verbs inflect not only for tense, mood, and subject person and number, but also according to certain features of the direct object. In the bulk of the cases where this phenomenon is relevant, the verb inflects differently according to whether or not the DO is a "definite" NP. In the latter case, the verb draws from a set of inflective endings collectively called the Objective Conjugation, hereafter referred to as OC. Various theories for what triggers OC endings, as well as theories on Object pro-drop licenced by said endings, will be described. The article will also explore the degree to which the OC morphology fuses with subject-inflective endings on the verb. Throughout, prominent theories will be peppered with the author's own opinionated viewpoints. I sincerely hope that, after reading this article, the reader will get a flavor for why this is such an interesting set of questions for linguists studying Hungarian, and for linguistics in general.

Basic Phenomena:

Here are some examples of how Hungarian selects different conjugations. First of all, if the DO is a DP headed by an indefinite article, OC is not triggered, as in (1), and if it is headed by a definite article, OC is triggered, as in (2):

(1a) Én látok egy hídf-
     I see one bridge-
     "I (can) see a bridge."

(b) Én látok hídak-
     I see bridges.

\* Except in possessed NPs, but I'll explain that later.

\* A note on my notation that I will use throughout this paper:
In my glosses, the ending on a verb will often be demarcated with a number (1, 2, or 3) and a letter (s or p). The number refers to person (first, second, or third respectively), and the letter to number (s=singular, p=plural). "acc"=accusative, "dat"=dative case, "perf"=perfective aspect, "pt"=past tense...

Also, I consistently use Hungarian spelling (rather than IPA) in the example sentences. This should cause no great problems, as most Hungarian letters correspond very closely to IPA anyway, or standard umlauting conventions. Exceptions include:

\* sz = [s], s=[f], a=[a], a=[e], e=[e], cs=[t]

The diacritic: above a vowel indicates the vowel is double-length, but the same quality except in the cases of 'a' and 'e'. Also, ő is long ő, and ű is long ű. 'y' after a consonant palatalizes it. 'j' after a consonant palatalizes it if possible -- otherwise, it is equivalent to IPA [c]. Geminated consonants are written double. So nn is long n. Also, nny is long ny, ggy is long gy, etc.
(2a) Én lát-om a hida-t.
I see-1s.OC the bridge-acc.
"I (can) see the bridge."
(2b) Én lát-om a hida-k-ot.
I see-1s.OC the bridge-p-acc.
"I (can) see the bridges."
(2c) Én lát-om Péter-t.
I see-1s.OC Peter-acc.
"I (can) see Peter."

Also, DO DPs headed by demonstratives ("this", "that", "these", "those"), as well as bare demonstratives, trigger OC:

(3a) Az-t a háza-t meg-ve-ttük.
that-acc the house-acc perf.-buy-pt.1p.OC.
"We bought that house."
(b) Az-t meg-ve-ttük.
"We bought that one."

(4a) Ez-ek-et a fá-k-ot lát-od?
this-p-acc the tree-p-acc see-2s.OC
"Do/Can you (sg.) see these trees?"
(b) Ez-ek-et lát-od?
"Do/Can you (sg.) see these (ones)?"

When the DO is a QP, only when it is headed by a universal quantifier will it trigger OC endings.

(5) Esz-nek néhány / sok / négy deka krumpli-t.
eat-3p some / many / four dekagram potato-ace.
"They’re eating (some / a lot of / four dekagrams of ) potatoes."

(6) Nem ért-ek semmi-t.
not understand-1s nothing-acc.
"I don’t understand anything."

(7) Minden-t meg-magyaráz-ta.
everything-acc perf.-explain-pt.3s.OC.
"(S)he explained everything. / (S)he explained all of it."
Finally (at least, finally in the simply explainable cases), no verb takes OC when being used intransitively:

(9) János alszik.
  John sleep-3s.
  “John is sleeping. / John sleeps.”

(10) A kutya nem jól lát-
  the dog not well see-3s
  “Dogs don’t /can’t see well.”

The distribution in sentences (1)-(8) fits very well with modern theories (as in The Representation of (In)definites) about what constitutes a definite or indefinite NP semantically. Sentences (9) and (10) provide motivation for seeing OC as an “extra” morpheme, rather than as one of two underlyingly equally likely choices of conjugations. Thus, if these were our only data, we could assume that the phenomena which trigger the OC in Hungarian are both semantically and syntactically transparent. We would say that it is semantically transparent, in that an object which is a definite NP triggers OC on the verb, and an object which is not a definite NP (or a null object altogether) fails to. We would say that it is syntactically transparent, in that NPs headed by definite articles, universal quantifiers, and demonstratives are exactly the NPs which trigger OC, while all others fail to.

However, there are some cases which do not neatly fit the mold. The behavior of OC with regard to pronominal subjects and objects, for instance, would cause one to doubt that this is a purely semantically driven phenomenon. And a look at DO pro in Hungarian will cast doubt on the idea that it’s a purely syntactic phenomenon. Building on some other work, we shall look at these and other phenomena as well, and develop a more sophisticated theory, which hopefully will account for the morphosyntactic and semantic distribution of OC.
Failure of certain pronouns to trigger definite-marking

The first problem with the idea that OC is triggered semantically by "definiteness" of the direct object (i.e. the state of affairs where both participants in the discourse are pragmatically expected to know the item referred to) is the fact that although third-person pronouns trigger OC, first- and second-person pronouns fail to, even though they clearly satisfy the traditional criteria of what definiteness is:

(i =f21)³ Péter lát engem-et / mink-et / téged / titek-et.
Peter sees me-acc / us-acc / you.sg.acc / you.pl-acc.
"Peter sees me / us / you."

(ii) a. Te lát-sz (/ Ti lát-tok) engem-et / mink-et.
You.sg see-2s (/ You.pl see-2p) me-acc / us-acc.
"You see me / us."

b. Te lát-od (/ Ti lát-játok) ö-t / Péter-t.
You.sg see-2s.OC (/ You.pl see-2p.OC) it-acc. / Péter-acc.
"You see it / Péter."

Another problem is the fact that the triggering of OC on the verb can depend on whether the direct object is marked for possession (a distinction the examples should make clear). Specifically, when the DO is so marked, the verb takes OC, even if the DO isn't semantically definite.

(iii =f16a) János lát egy kutyá-t.
John sees a dog-acc
"John sees a dog."

(iv) János lát-ja egy kutyá-já-t.
John sees-OC a dog-3s.poss-acc.
"John sees one of his dogs."

(v) János lát-ja a kutyá-já-t.
John sees-OC the dog-3s.poss-acc.
"John sees his dog."

There are two more anomalies sometimes associated with the OC question in

³The examples (i-vii) here are repeated later in the section, and given lower-case Roman numerals here so as not to create conflicting numbering systems. The examples in the footnotes will be given upper-case Roman numerals corresponding to the number of the footnote.
conjugating Hungarian verbs. First, there is the so-called "I-you" construction. Namely, there is a morpheme that the verb gets inflected with only when the subject is first-person singular and the DO is second-person (singular or plural):

\[(vi) \quad a. \quad \text{Én lát-om} \quad \ddot{a}-t \quad / \ddot{a}-k-et. \]
\[\text{I see-1s.OC 3s-acc / 3s-pl-acc.} \]
\[\text{"I see her / him / them."} \]

\[b. \quad \text{Én lát-lak} \quad \text{tégéd} \quad / \text{titek-et.} \]
\[\text{I see-1s.2obj you.sing.acc / you.pl.acc-acc.} \]
\[\text{"I see you."} \]

Second is the treatment of anaphors. Intuitively it is difficult to see whether reflexives should be seen as definite or indefinite. On the one hand, they stand for elements of the discourse of which clearly both participants are expected to be aware. On the other hand, anaphors always have their referents within the same clause.

\[(vii = f^{11}c) \quad \text{Mari lát-ja} \quad \text{magá-t.} \]
\[\text{Mari see-3s.OC self.3s-acc} \]
\[\text{"Mari sees herself."} \]

In this section, we shall investigate the approaches of the linguists Daniel Abondolo and Donka Farkas on these matters and attempt to come up with a workable synthesis and fine-tuning of their ideas.

**Abondolo's approach:**

Daniel Abondolo crucially connects the pronoun question to the question of the "I-you" construction. He does this by conceiving the OC phenomenon as an issue of explicit vs. implicit object marking (rather than the traditional [-def] vs. [+def] feature morphology, or subjective vs. objective conjugations, or Farkas' "OC", which I use here). It is questionable whether his theory could be extended to many other languages, but it is interesting and seems to account well for most (though not all) of the phenomena.

Let us examine Abondolo's model. According to Abondolo, "the expression of object person in Hungarian begs a gradient analysis, according to which the category of person is arrayed along a scale ranging from least ambiguous (first person, ...) to most

\footnote{Sometimes anaphors needn't be in the same clause as their referents, but they must always in some sense be "near" them; this varies from language to language, and is generally agreed to be a messy issue. Indeed, they seem to have been a large part of the inspiration for G-B theory and control theory. See, for instance, Chomsky's 1981 monograph: *Lectures on Government and Binding*}
ambiguous (third person,...)." (p.88) In other words, second-person is less ambiguous than third, but more than first, so we get his "gamut of increasing ambiguity" represented in the form of a gradient, or as concentric circles. The closer we are to the center (or the left edge) in this model, the less ambiguous the referent.

1 -----> 2 -----> 3 -----> 3''

Why would it be said that first-person is less "semantically ambiguous" than second-, or second-person than third-person definite? It's less about ambiguity and more about context dependence. For instance, the meaning of the words "I" and "me" depend only on the speaker, whereas "you" depends both on the speaker and the listener, and the meaning of the phrase "the rabbit" depends on the whole conversational context. On the other hand, the meaning of the phrase "a rabbit" is completely independent of any context except sharing of a common language.

His explanation for the phenomena rests on the claim that selection of verbal conjugation (non-OC or OC) in Hungarian sentences that "do not explicitly indicate the existence of an object" is a matter of determining the relationship of the subject to the direct object in terms of relative position in the person gradient (i.e. in terms of relative levels of "ambiguity").

Specifically, non-OC corresponds to those cases where the object is either unspecified or is earlier in the gradient than the subject. We call this "implicit object marking" because persons earlier in the gradient, being less semantically ambiguous (i.e. less dependent on context) than those later, trigger more "implicit" marking than those later. Likewise, OC corresponds to cases where the DO is specified and is later in the gradient than the subject, so that it depends on context more than the subject does, thus triggering more "explicit" marking than the subject. Furthermore, the reason we get a special form in the case of a 1st person singular subject with a 2nd person object, is because truly explicit object marking should indicate what sort of object it points at. Since both 2nd person and 3rd person are further forward in the gradient than 1st person, there should be a separate form for each. This theory correctly predicts examples 11-12 below.

(11 =f21) Peter lát engem-et / mink-et / téged / titek-et.  
  Peter sees me-acc / us-acc / you.sg.acc / you.pl.acc.  
  'Peter sees me / us / you.'

3 The symbol: 3' refers to a third person object. The argument goes: since there are so many third-persons, some languages (such as Hungarian) differentiate roles of third person subject and third-person object more than they do for 1st-person or 2nd-person.

4 which I shall denote as "1s.2obj", as in (13b)
(12) a. Te lát-sz (/ Ti lát-tok) engem-et / mink-et.
   You.sg see-2s (/ You.pl see-2p) me-acc / us-acc.
   'You see me / us.'

   b. Te lát-od (/ Ti lát-játok) űt / Péter-t.
   You.sg see-2s.OC (/ You.pl see-2p.OC) it-acc. / Péter-acc.
   "You see it / Péter."

   c. Én lát-om (/ Mi lát-juk) űt / Péter-t.
   I see-1s.OC (/ We see-1p.OC) it-acc. / Péter-acc.
   "I see (/ We see) it / Péter."

This theory can be strengthened a bit by changing the gradient diagram to look like
the following one:

(13) a. 3[-def] ----> 1 ----> 2 ----> 3[subj, +def] ----> 3'[DO, +def]

This is fully supported semantically (except, perhaps, for the last arrow, but that’s
forgivable), because, for example, it predicts correctly that the following four phrases
are sorted in order from least context-dependent to most context-dependent.

   b. "a rabbit" ----> "I" / "me" ----> "you" ----> "the rabbit"

Consider: The phrase "a rabbit" depends for its meaning neither on the identity of the
participants in the conversation nor on the content of the conversation, "I" depends on
the identity of the speaker, "you" depends on the identities of both participants, and
"the rabbit" depends on the whole conversational context. Given this improved
gradient, then, we can explain almost all the foregoing data in the following terms: if
going from subject to object follows the direction of the gradient (13a), OC
morphology gets triggered. Otherwise, it doesn’t.

Far as the gradient analysis takes us, there are some weaknesses in it. For one thing,
Abondolo does not attempt to account for the fact that possessed objects, even if
preceded by the indefinite article, trigger OC.

(14 =f16a) János lát egy kutyá-t.
   John sees a dog-acc
   'John sees a dog.'
Consider (14) and (15) from Farkas. The concept of "a friend of his" seems to me no more explicit in Abondolo's sense than that of "a dog", and yet the former triggers OC marking on the verb, while the latter does not.

Abondolo also does not give us any prediction for whether or not reflexive objects should trigger OC, because the gradient analysis only provides for those cases where the subject and object are of different persons, or both 3rd person (see footnote #5).

Also, even in what he does supposedly account for, there are some shaky issues. According to the gradient analysis, we would expect the coupling of a plural first person subject with a second person object to give us a different form than the thirteen we already have, just as the coupling of a singular first person subject with a second person object gave us a different form than the twelve we already had. But there are only thirteen forms, as it turns out. According to Abondolo (p.93) "This hiatus is a reflection of the asymmetry inherent in the deixis of person: first person plural is not a sum of first persons singular, plurality of speaker being in fact impossible."

It is true that in my context-dependence interpretation of Abondolo's theory, "we" works differently than "I". The former depends on the identities of both the speaker and the speaker's comrades, whereas the latter depends only on the identity of the speaker.

However, even if first-person plural subjects don't act quite like singular ones, we would at least expect the same form as that which a first-person plural subject triggers with a definite third-person object, since the gradient predicts that the object marking
would be "explicit". Instead, as shown in (17b), we get the implicit, or standard non-OC, marking form.

(17) b. Mi lát-unk téged / titek-et.
    We see-1p you.sing.acc / you.pl-acc.
    "We see you."

And this holds true as well in the other three morphologically marked tense/moods in Hungarian, as shown in 17d (in contrast to (17c)).

(17) c. Én lát-(ná / ja / ta)-lak téged / titek-et.
    I see-(cond / subj / past)-1s.2obj you.sg.acc / you.pl-acc.
    "I (would / should / did) see you."

(17) d. Mi lát-(ná-nk / j-unk / t-unk) téged / titek-et.
    We see-(cond-1p / subj-1p / past-1p) you.sg.acc / you.pl-acc.
    "We (would / should / did) see you."

Farkas’ model:

Farkas, on the other hand, argues for an underspecified feature, [def], in Hungarian morphology, the marking of which may be blocked by certain conditions on the features of the DO. Her theory accounts for most of the phenomena, and moreover uses terms in common use in modern linguistic theory. However, her model does not

 Accordingly, we may want to adjust the gradient once more, to look like this:

(VIII) 3[-def] ----> 1s ----> 2 ----> 1pi ----> 3[subj, +det] ----> 3[DO, +def] :

After all, it could be argued that "we" involves just as much (or, crucially, more) context as (or than) "you", as follows: The word "you" involves the identities of a speaker and of a set of (1 or more) listeners. The word "we" or "us", when used in a way so as to include the listener(s), involves the identities of exactly the people "you" involves: the speaker and the listener(s). However, when "we" or "us" is used so as to exclude the listener(s) (for instance, when used in narrative speech, as in the sentence "We went to the park today."). it involves the identities of the speaker and also a referent of the discourse (i.e. the same sort of people 3rd person definite NPs refer to). Thus, in the first case, it is just like second-person, whereas in the second case, it is just like third-person-definite. So semantically it makes some sense to put it between second-person and third-person-definite on the gradient diagram.

Of course, this begs the question of what to do with sentences where the subject is first-person-singular and the DO first-person-plural. Such sentences, due to the nature of human discourse, are few and far-between, and as I do not speak Hungarian natively, I don’t know whether these sentences trigger OC or not. The gradient exhibited in this footnote predicts that they would, but my suspicion is that they don’t. Also, this would be incompatible with other theories (e.g. Farkas'), which seem to have more universal scope. Accordingly, I will not take the view that (VIII) is correct, although I shall leave it as an interesting possibility for other linguists to wrangle over.
attempt to explain why certain features should take precedence over others.

She suggests that in Hungarian, “person features are semantically driven nominal features whose content concerns the relation of the referent of the NP to the participants in the speech act.” (p. 547). Specifically, two features are relevant in determining person: [prt], or “participant”, and [sp], or “speaker”. First- and second-persons are “participants”, whereas third-person is not. First-person is “speaker”, whereas second-person is not. Schematically (her diagram (=f22)), we have

```
/+sp/ (1st person)
/    
/+prt/ 
/        
/           
/               / 
/               / [sp] (2nd person)
/               
/               [-sp] (3rd person)
```

There are many good reasons why it is reasonable to assume that Hungarian regards first- and second-persons as the components of a natural class in opposition to third-person.°

She says, “the OC endings are triggered by DOs that have the feature [def].”, which she says has semantic content: “—namely, reference to a ‘familiar’ discourse referent, in the sense of Heim (1982).”

°For instance, in simple sentences composed of subject, copular verb, and a predicate NP, the verb is phonetically null only when the subject is third-person:

```
(IX)  a. (En) tanár vagy-ok.
      (I) teacher be-1s
      “I’m a teacher.”

b. (Te) tanár vagy.
      (You) teacher be-2s
      “You’re a teacher.”

c. Ő tanár.
      he/she teacher
      “(S)he’s a teacher.”
```

Further reasons to assume this are contained in Farkas’ treatment of OC.
However, she notes that this does not fully account for all instances in which OC is used, nor for all instances where it is not used. For those cases which call for OC but are not semantically [def] in this sense, she posits the following FCR (“Feature Cooccurrence Restriction”) rules to deal with them:

\[(18 = f24)\]

\[\begin{align*}
  &a. \ [+s(\text{sentential})] \to [\text{def}] \\
  &b. \ [\text{refl(exive)}] \to [\text{def}] \\
  &c. \ [\text{poss(essive)}] \to [\text{def}]
\end{align*}\]

To account for the fact that verbs with DOs which are [+prt] do not take OC endings, she posits the following FCR rule:

\[(19 = f23) \quad [+\text{prt}] \to [\text{def}]\]

She calls this an “inherent” rule, because the referent of a first- or second-person pronoun is a participant in the speech act, and thus semantically definite. At first, one finds this sort of thing rather fishy. Two questions jump to mind:

(20) If rule (19) is inherently true by semantic content of the feature [+prt], why go to the trouble of stating it as a separate rule?

(21) Okay, so (19) is true. We know it inherently, and we stated it explicitly for some reason too, so it’s doubly true. So why don’t first- and second-person pronominal DOs trigger OC endings?

Farkas offers three explanations, each based on a different theoretical approach. (22a) is the generalization, (22b) is what we say in a “rule-ordering” approach, and (22c) is what we say in a “counterpart” approach, to explain what is happening.

(22) a. “the feature introduced by (19) is inert with respect to OC”
   b. “[+prt] pronouns are underspecified for [def]” and (19) is ordered after the rule OC.
   c. “the PF introduced by (19) does not have an SF counterpart.”

In general, the ability of the semantic rule that marks [+prt] as [def] is commandeered by (19), which is a morphological rule and which, thus, is susceptible to further manipulations and stipulations.

Farkas succeeds in using modern linguistic theory to explain most of the morphosyntactic behavior of Hungarian OC. However, she doesn’t touch the special
I would like to propose a synthesis, conflating Farkas' and Abondolo's work on this matter.

**A Synthesis:**

I will adopt Farkas' treatment of [def] as an underspecified feature, as well as her featurewise treatment of person, and I will assume that the decision of what verbs get OC is based almost purely on the semantics of the subject and direct object of the verb, as does Abondolo. However, I have some reservations with both of their theories.

First of all, I disagree with Abondolo's hypothesis that the "I-you" conjugation is anything less than an exceptional case, morphosyntactically speaking. Nowhere else in the paradigm of where to use OC is number a relevant feature (but see footnote #8 above)... Much as exceptions are odious, we must accept this one, assuming that my hunch in the aforementioned footnote is correct.

Now, Fárkas only gives an outline-like explanation for why we should expect [poss] to trigger OC. Furthermore, she does not give any semantic reason for why [refl] or [+s] should trigger the same morphology that [def] triggers. There's nothing inherently wrong with this omission; sometimes morphology is triggered by phenomena that have no apparent link to each other at all. However, I think there's something more coherent going on than is being accounted for. In my explanation, we don't even need to deal with the FCR (19).

According to Heim (1982), apparently, Farkas claims that reflexive pronouns should not be semantically [def]. I have not read Heim's article, but I suspect that different languages have different notions about what is and is not a "definite" NP. I would propose that an NP in Hungarian should be considered [def] semantically if its referent is one that the listener should be able to determine exactly, straight from the semantics of any clause containing the NP as a DO. If we look at things this way, then reflexives are inherently definite, since any clause that contains a DO reflexive must have a...

\*Furthermore, it still seems suspect to me to postulate a rule like (18), claiming that one domain of the grammar (morphology) can wrest a feature from a distant other domain (semantics) in some but not all cases.
subject whose referent is the same as the reflexive's referent. Subordinate clauses are also definite, if we consider them NPs."

As for possessives, Farkas explains that "the semantic content of [def] is extended to cover not only familiar reference but also reference to an entity connected to a familiar referent. The familiar referent is the possessor and the connection is the possessive relation. This type of definiteness will be called indirect." Instead, I would propose a sort of feature percolation, as illustrated below:

(22) NP - <possessor ending>
     \   II
     \   II
      [def]

Let us assume that OC is triggered according to two parameters -- [person] and [def], with the rule schematically described below:

```
     [def] non-OC
     Spers=DOpers CC
     DO=[+prt] non-OC
```

interpret the tree above as follows: Follow the left branch down from the top as far as you can, and if at any point you can't, pick the nearest preceding right-branch. If you aren't convinced that sentential objects are definite, place a node that says "O ≠ [+s]" somewhere on the left branch between [def] and "non-OC". But it's much more elegant if we just assume sentential objects are OC.


It may seem suspect to assume that endings that mark possessiveness attach to the NP rather than the noun, but it has been argued (Napoli, 1996) that case endings do just that, which prefigures a pseudo-incorporation type theory, but the cited source argues for just that. It's weird and interesting, and it's the only thing that allows one to argue that Hungarian really has free word order.

"Farkas (1987) postulates the following rule, which closely resembles the implications of the above tree:

((16) in Farkas, p.200) V gets OC in the environment: Object is [refl] or both [-prt] and [def]. The crucial difference between her approach and mine is that she uses [refl] as a separate feature, rather than treating it, as I do, as just an NP with the same person value as the subject."
At this point, the reader might protest that this is no simpler than Farkas' model. I agree; it isn't any simpler. The phenomenon is complex, and thus requires a certain amount of complexity in any coherent explanation of it. But I think my explanation is more algorithmically elegant, and trees are easy to follow, like neural pathways... If we assume that OC is dependent on characteristics of person in subject and DO, and definiteness in DO, we get a tree algorithm with just two branching points. The reason for the proposed extra node in the last paragraph, moreover, is that sentential objects, unlike NPs, don't really have a person feature. Either that, or they are third-person by default (since they certainly are not [+prt], going by Farkas' feature model of person, so from an underspecification viewpoint, they would likely behave the same way as NPs which are [-prt]), in which case the tree holds as depicted.

In words, we may summarize the tree by saying that a verb becomes a candidate for OC only when, either the Subject and DO agree in person, or the DO is [-prt] (or [+s]). Later, those candidates whose objects are [def] get OC endings. Thus, the subject-object person-person relationship is what licenses OC in object-[def] environments.
Direct Object pro

Farkas (1987) is an excellent article about DO pro in Hungarian, in relation to various recent theories of pro. The author proposes the following "diagnostic properties of pro" (which I have simplified a bit), and shows that Hungarian exhibits them with DOs:

(23 =f2) Diagnostic properties of pro
a. lexically unrestricted occurrence
b. possible controller of infinitives
c. governed position
d. may be coreferential with a non-topic NP in the matrix clause
e. Case-marked
f. not coindexed with an element in the specifier position of C
  g. cannot replace an overt reflexive pronoun

Then, taking into account how Hungarian DO pro works, she compares various other people's theories about pro, arguing for some theories and against others. As a crucial part of her argument, however, she claims that OC marking (which she seems to assume is a prerequisite for DO pro in Hungarian) is not entirely based on features of person and definiteness. Thus, she says, Hungarian is a counterexample to Huang's application of Taraldsen's generalization, that "pro must exhibit 'strong' agreement with the V. Agreement is generally supposed to be 'strong' if it includes at least agreement in person." (p.194, "cf. Taraldsen 1978, Chomsky 1981, among many others"). I believe that this is partially because Farkas only considers those cases where a DO pro is coupled with OC on the verb. I'd like to investigate the possibility that there are cases of DO pro in Hungarian where the verb does not have OC. After looking at those cases, we shall hopefully see that person features are at least partially recoverable, so that Hungarian isn't a counterexample after all.

First we will show, using Farkas' exact method of proof, that certain DO ECs in Hungarian (including those that Farkas considers as well as others) exhibit the properties in (23), and thus should be considered pro. To prove (23a) about null DOs in Hungarian, she simply explains that "the DO of any transitive V may be null." in Hungarian. To prove (23b), she uses the following sentence (24a).

---

14 In terms of the theory of Chomsky (1986).

15 Except for once, where she does mention the "l-you" conjugation, and (correctly, in my opinion) blows it off as a special case, rather than as a phenomenon regular enough to generate a theory from.

16 The four examples I use in this capacity (parts (b) of sentences 24-27) I was unfortunately not able to check with a native speaker of Hungarian, nor could I find similar sentences, nor any indications that such constructions were ungrammatical, from any of my experience living in Hungary for four months or my subsequent reading. I will assume they are grammatical, because intuitively it seems they should be. However, if they are not, my argument about DO pro clearly does not hold true.
In both sentences in (24), the EC controls the infinitive usz-ni, "to swim", so (23b) is proved.

Next, consider the following examples:

(25a) János tud-ja, hogy Mari szeret-e, j', k
(=f1a) János know-3s.OC that Mari love-3s.OC him.
"János knows that Mari loves him;"

(b) Én tud-om, hogy Mari szeret-e, j', k
I know-1s.OC that Mari loves-3s (me).
"I know that Mari loves (me)" or "I know that Mari is capable of love."

It is clear from (25) that the ECs shown are in governed and case-marked position (governed by the verb szeret-, "to love"; case-marked accusative), as well as in the sentences in (24) and (26), so (23c) and (23e) are proved.

(26a) János megmond-t-a Mari-nak, hogy Gábi utal-ja, e, i, j, *k
(=f1b) János tell-3s.OC Mari.dat., that Gábi hate-3s him/her.
"János told Mari that Gábi hates him/her;"

(b) Én megmond-om neked, hogy Gábi utal-e, i, j, *k
I tell-1s.OC you.dat., that Gábi hate-3s me/you.
"I'm telling you that Gábi hates me / you;"

In the sentences (26), since the EC in each case can be coreferential with either the NP subscripted with i, or the NP subscripted with j, and since there can be at most one
topic NP per clause, it follows that the EC may be coreferential with a non-topic NP in the matrix clause, proving (23d). As for (23f), it is clear from the examples given that none of these ECs are specifiers of C' (i.e. traces of wh-phrases). Farkas shows that (23g) is true of sentences where the verb takes OC, and since the only sentences where the DO is reflexive trigger OC on the verb, (23g) holds vacuously for sentences where the verb does not take OC. Therefore, it is the case that the ECs in sentences like (24b), (25b) and (26b) exhibit Farkas' "diagnostic properties of pro" just as surely as, as Farkas showed, sentences like (24a), (25a), and (26a) exhibit them. More generally, we can say that, barring traces, implicit arguments, and PRO, phonetically null DOs in Hungarian are instances of pro, whether or not the verb exhibits OC morphology. Let us investigate this, then. As a variant on (26), consider (27):

(27a) \( \text{Én} \text{ megmond-ta-m Ő Peter-nek} \text{, hogy \( \text{Gábi} \text{k meg-hoz-ott} \text{ e} \text{i, j, k} \text{.} \) \)} \\
I tell-pt-1s.OC Peter-dat., that Gábi perf.-bring-pt me\text{.} \\
"I told Peterj that Gabi\text{k} brought me\text{.}"

(27b) \( \text{Én} \text{ megmond-ta-m Ő Peter-nek} \text{, hogy \( \text{Gábi} \text{k utal-ja \text{ e} \text{i, j, k} \text{.} \) \)} \\
I tell-pt-1s.OC Peter-dat., that Gábi hate-OC him. \\
"I told Peterj that Gabi\text{k} hates him\text{.}"

There certainly seems to be a large degree of recoverability here. As discussed in the section on person marking, OC marking occurs exactly in those cases where the object is definite where either the object is third-person, or the subject and object are the same person. Since any kind of pro is required to refer to a referent already in the discourse, any pro DO will have the inherent feature [def]. Thus, if there is a DO pro, its verb is OC if and only if it is either of the same person as the subject or in third-person. So we have the following paradigm for DO pro: Reflexives are implicitly excluded, of course.

<table>
<thead>
<tr>
<th>Su person</th>
<th>DO person</th>
<th>OC?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>no</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>yes</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>no</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>yes</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>no</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>no</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>yes</td>
</tr>
</tbody>
</table>

\(^{17}\) When the subject person is singular, then yes, this is the infamous "I-you" construction. But it's still not traditionally considered OC, and that tradition clearly suits our purposes here.
So the antecedents of both Su pro and DO pro are always marked for person by the verb. The verb is marked for person and number of the subject. Then if the DO is an empty category (and not a trace or a PRO), a first-person subject dictates 2nd person object if no OC on the verb, 3rd person object if there’s an OC on the verb. A second person subject dictates 1st person object if no OC on the verb, and 3rd person object if there’s an OC on the verb. A third-person subject dictates [+prt] object if no OC on the verb, and [-prt]=3rd person object if there’s an OC on the verb. The inflection on the verb will recover the person and number of the Subject, and at least one, if not both, features of the person of the DO, if either the Subject or DO (or both) is an EC in the sentence. Tidy.
References


---

*some of these are copied straight from Farkas’ bibliographies at the end of the two articles of hers I’ve cited here*