

The lack of full pro drop as a consequence of featural overspecification

Olaf Koeneman and Hedde Zeijlstra

1 Introduction

Some languages allow the argumental subject of the sentence to remain unexpressed, whereas other languages do not. Italian is a so-called *pro drop* language, whereas English is not:

- (1) a. Gianni ha detto che ha telefonato Italian
 Gianni has said that has.3SG telephoned
 ‘Gianni said that he called’
 b. *John said that telephoned English

Setting apart radical pro drop languages, in which all arguments can be dropped, as a separate phenomenon (cf. Neeleman & Szendrői 2007 for discussion and references), one can establish a correlation between the possibility of having null subjects and richness of agreement. Italian has a rich agreement paradigm, whereas English does not, so that only in the former language can a missing subject be reconstructed. From a naïve functional perspective, however, it is then not immediately obvious why a null subject is not licensed in English 3SG contexts: the *-s* affix is as unique to the English paradigm as the Italian *-a* affix is to the Italian one. A popular solution is to refer to the entire paradigm: Italian is rich overall, and English is poor overall, so that we can hinge the licensing of null subjects onto a parameter that refers to the paradigm. This may even help to understand why rich Germanic languages, such as Icelandic and Standard German, lack null argumental subjects: their paradigms contains a syncretic form and this might make the paradigm non-rich (cf. Koeneman 2000; Tamburelli 2006).

	Italian	English	Icelandic	St. German	Romanian
1SG	-o	-∅	-i	-e	-0
2SG	-i	-∅	-ir	-st	-i
3SG	-a	-∅	-ir	-t	-ă
1PL	-iamo	-∅	-jum	-en	-ăm
2PL	-ate	-∅	-ið	-t	-ți
3PL	-ano	-∅	-ja	-en	-ă

Table 1: Agreement paradigms in Italian (-are conjugation), English, Icelandic, Standard German and Romanian (1st conjugation)

Such a paradigmatic approach, however, runs into several problems. First of all, paradigms are epiphenomenal, a handy tool for the linguist but not a construct that native speakers make reference to in their grammar (cf. Bobaljik 2003). Second, Romanian also has a syncretism in its paradigm (for instance in the 3SG and 3PL cells in the first conjugation). Yet it is a full-fledged null subject language. Third, the existence of partial pro drop in languages like Standard Finnish, Hebrew and certain Germanic dialects, where null subjects are only allowed in certain person/number contexts, strongly suggests that an all-or-nothing parameter approach misses the mark. Given these problems, one could of course abandon the paradigmatic approach in favor of a contextual one, where the possibility of a null subject must be determined for each specific context. But that begs the question why the 3SG context in English does not license a null subject, and this problem of course aggravates if one wants to understand the lack of pro drop in richer Germanic languages, like Icelandic and Standard German.

To conclude, a paradigmatic approach to pro drop *undergenerates* (as it does not expect partial

pro drop) whereas a contextual approach *overgenerates* (as it expects too much (partial) pro drop). The question is how to get out of this conundrum.

We argue that it is possible to derive paradigmatic effects in a contextual approach. The key to understanding the lack of pro drop in Germanic languages is the following empirical generalization. If you look at present and past agreement paradigms, then the form that appears in the 3SG cell of the present tense never returns in the 3SG past tense.¹ Strikingly, as we will see later on, this is not the case in those languages that exhibit classical pro-drop, such as Italian, Spanish and Romanian; here, the 3SG form in the present tense reappears in the past tense.

	Icelandic		Standard German		English	
	Present	Past	Present	Past	Present	Past
1SG	-i	-ð-i	-e	-te	-∅	-ed
2SG	-ir	-ð-ir	-st	-te-st	-∅	-ed
3SG	-ir	-ð-i/*ð-ir	-t	-te/*-te-t	-s	-ed/*-eds
1PL	-jum	-ð-um	-en	-te-n	-∅	-ed
2PL	-ið	-ð-uð	-t	-te-t	-∅	-ed
3PL	-ja	-ð-u	-en	-te-n	-∅	-ed

Table 2: Present and past tense paradigms in Icelandic, Standard German and English

The disappearance of this 3SG form in the past must be captured by the morphological analyses of these paradigms. As we will show in section 2, the consequence of this will be that the tense and agreement paradigms in the Germanic languages are qualitatively less transparent than Romance languages like Spanish and Italian in their encoding of tense and agreement. The repercussions of this for pro drop will be discussed in section 3, where we argue that qualitative transparency, here the fact that tense and agreement are morphemically distinct markers, is a necessary condition for pro drop. In section 4, we show how our proposal extends to partial pro drop. Section 5 concludes.

2 Morphological analyses of Germanic and Romance paradigms

2.1 Icelandic and Standard German

At first glance, Icelandic looks like a language in which tense and agreement can be straightforwardly distinguished. We can take $-\ð$ to be the past tense marker which is followed by agreement markers. The spell-out rules for tense would then look as in (2), where a null marker spells out unmarked tense (= present tense):

- (2) Tense
- | | | |
|--------------|-------------------|-----------|
| $-\ð$ | $\langle \rangle$ | [T: past] |
| $-\emptyset$ | $\langle \rangle$ | [T:] |

To capture the syncretism in the 2SG and 3SG slots in the present tense, we can adopt the following agreement spell-out rules:

- (3) Tense
- | | | |
|-------|-------------------|----------------------|
| $-i$ | $\langle \rangle$ | [u ϕ : speaker] |
| $-ir$ | $\langle \rangle$ | [u ϕ :] |

In the slipstream of Benveniste (1971), substantial evidence has been put forward for the assumption that the 3rd person is a non-person, expressing no feature values (cf. Harley & Ritter 2002; Preminger

¹ As far as we can tell, this generalization is exceptionless. Dutch and its dialects, as well as Faroese, have subject agreement and their 2/3SG $-t$ and $-ir$ forms, respectively, do not return in the past tense. Ölvödalen Swedish has a marker for singular number, $-är$, that does not return in the past tense (Garbacz 2010:45).

2014). If so, the form appearing in the 3SG slot must count as the elsewhere form. Therefore, *-ir* is the agreement elsewhere. The question is how to account for the fact that this form does not reappear in the 3SG past tense slot, where we find *-i* instead. In an attempt to maintain the rules in (3), we could resort to an impoverishment rule. The appearance of a different form is the consequence of a feature value that has been erased so that *-ir* can no longer be inserted. The problem, however, is that *-ir* is the elsewhere form. Hence, there is no feature value that can be impoverished. We could alternatively abandon the assumption that *-ir* is the elsewhere and assume that *-i* functions that way, but that does not help. We would need the spell-out rules in (4a) and the impoverishment rule in (4b):

- (4) a. *-ir* \diamond [u ϕ : non-speaker]
 -i \diamond [u ϕ :]
 b. [u ϕ : non-speaker] \rightarrow [u ϕ :] / __[u ϕ : non-addressee], [T: past]

It derives the facts but the rule in (4b) is suspicious. In order to ensure that 2SG *-ir* does not change to *-i* in the past too, the impoverishment rule needs a context-feature ([u ϕ : non-addressee]) that is not mentioned in the spell-out rules in (4a). Even worse, [u ϕ : non-addressee] could not be generated in a syntactic agreement slot as that position is already taken by [u ϕ : non-speaker].

Resorting to impoverishment, then, is not a possible way to account for the agreement alternation between present and past. We could alternatively make use of a context-sensitive rule:

- (5) *-i* \diamond [u ϕ :] / [T: past]

This rule overgenerates, as we now expect *-i* in 2SG past contexts too. To circumvent this, we must postulate two *-ir* forms, as in (6):

- (6) *-ir* \diamond [u ϕ : addressee]
 -ir \diamond [u ϕ :]

This derives the facts but at the cost of creating two homonyms. Since the analysis must postulate two *-i* and two *-ir* forms, it fails to capture the syncretic patterns in the Icelandic paradigm.

A third way to capture the agreement alternation is to give up the idea that tense and agreement are separate morphemes. This would constitute an analysis with hybrid morphemes that express tense and agreement features at the same time. We could think of this morpheme as INFL. The spell-out rules then look as follows:

- (7) *Inflection (tense and agreement):*
- | | |
|---|--|
| <i>-i</i> \diamond [T:], [u ϕ : speaker] | <i>-ði</i> \diamond [T: past], [u ϕ :] |
| <i>-ir</i> \diamond [T:], [u ϕ :] | <i>-ðir</i> \diamond [T: past], [u ϕ : addressee] |
| <i>-jum</i> \diamond [T:], [u ϕ : speaker, plural] | <i>-ðum</i> \diamond [T: past], [u ϕ : speaker, plural] |
| <i>-ið</i> \diamond [T:], [u ϕ : addressee, plural] | <i>-ðuð</i> \diamond [T: past], [u ϕ : addressee, plural] |
| <i>-a</i> \diamond [T:], [u ϕ : plural] | <i>-ðu</i> \diamond [T: past], [u ϕ : plural] |

This mono-morphemic way of representing the Icelandic paradigm may look underanalyzed but it gives clear advantages. The 3SG form *-ir* can be maintained as the elsewhere, the default option, and the rules do not postulate any homonymous forms: all forms are clearly distinct. No impoverishment rule has to be postulated and no context-sensitive rules either. Note that a bi-morphemic analysis would already need three context-sensitive rules to capture the three agreement alternations in the plural (*-jum* vs. *-um*, *-ið* vs. *-uð* and *-a* vs. *-u*), so that a last advantage is that the mono-morphemic analysis also needs fewer rules (10 versus 12, not shown in detail here).

Similar issues apply to Standard German. Again, the problem to solve is an agreement alternation, the non-appearance of the 3SG present tense form in the past, where we find *-te* instead of *-tet*. (cf. Table 2). An impoverishment solution is unavailable for a 3SG context under the assumption that there are no features to impoverish in the elsewhere environment. It is technically possible, however, to induce a context-sensitive rule, which would look as in (8):

(8) $-\emptyset \quad \langle \rangle \quad [\text{u}\varphi:] / [\text{T}: \text{past}]$

What this rule does is introduce a second elsewhere form and this elsewhere form has to be phonetically empty. One may first of all wonder whether having a second elsewhere form does not defy the purpose of an elsewhere to begin with. More importantly, however, one may wonder whether the postulation of a null allomorph is an easy step to take in the acquisition process. It may be an unavoidable step for a child to take in the absence of an alternative analysis, but we have just seen for Icelandic that an alternative analysis in fact exists, namely the mono-morphemic one in (9):

(9) *Inflection*

$-e \quad \langle \rangle \quad [\text{T}:], [\text{u}\varphi: \text{speaker}]$	$-te \quad \langle \rangle \quad [\text{T}: \text{past}], [\text{u}\varphi:]$
$-st \quad \langle \rangle \quad [\text{T}:], [\text{u}\varphi: \text{addressee}]$	$-test \quad \langle \rangle \quad [\text{T}: \text{past}], [\text{u}\varphi: \text{addressee}]$
$-t \quad \langle \rangle \quad [\text{T}:], [\text{u}\varphi:]$	$-ten \quad \langle \rangle \quad [\text{T}: \text{past}], [\text{u}\varphi: \text{plural}]$
$-en \quad \langle \rangle \quad [\text{T}:], [\text{u}\varphi: \text{plural}]$	$-tet \quad \langle \rangle \quad [\text{T}: \text{past}], [\text{u}\varphi: \text{addressee, plural}]$
$-t \quad \langle \rangle \quad [\text{T}:], [\text{u}\varphi: \text{addressee, plural}]$	

What ties together Icelandic and Standard German, then, is that an analytical choice has to be made and that it depends on one's assumptions on parsimony if the choice falls out in favor of a mono- or bi-morphemic analysis. A similar choice arises for English. If one sticks to the assumption of the 3SG form as the elsewhere form, impoverishment or context-sensitivity must be induced to account for the disappearance of $-s$ in the past tense, and this creates similar issues. A mono-morphemic analysis would account for the facts without inducing impoverishment or context-sensitivity. It needs to postulate two null forms but so does a bi-morphemic analysis, as shown in (10).

(10) *Mono-morphemic analysis*

$-\emptyset \quad \langle \rangle \quad [\text{T}:], [\text{u}\varphi: \text{participant}]$
$-\emptyset \quad \langle \rangle \quad [\text{T}:], [\text{u}\varphi: \text{plural}]$
$-s \quad \langle \rangle \quad [\text{T}:], [\text{u}\varphi:]$
$-ed \quad \langle \rangle \quad [\text{T}: \text{past}], [\text{u}\varphi:]$

Bi-morphemic analysis

$-s \quad \langle \rangle \quad [\text{u}\varphi: \text{non-participant}]$
$-\emptyset \quad \langle \rangle \quad [\text{u}\varphi:]$
$-\emptyset \quad \langle \rangle \quad [\text{T}:]$
$-ed \quad \langle \rangle \quad [\text{T}: \text{past}],$ $[\text{u}\varphi: \text{non-participant}] \rightarrow [\text{u}\varphi:] / [\text{T}: \text{past}]$

In sum, we have pointed out that morphological analyses of Icelandic, Standard German and English need to capture the fact that the 3SG present tense form does not return in the past tense and that this requires a particular analytical choice in each case. In the next section we will show that the analyses for Italian and Spanish are more straightforward and that the mono-morphemic analysis therefore does not come into play.

2.2 Spanish and Italian

Spanish and Italian have a very transparent tense and agreement system. As can be established in Table 3, most agreement affixes appearing in the present tense come back in the imperfect:

	Spanish		Italian	
	Present	Past	Present	Past
1SG	-o	-ab-a	-o	-av-o
2SG	-as	-ab-as	-i	-av-i
3SG	-a	-ab-a	-a	-av-a
1PL	-amos	-áb-amos	-iamo	-av-amo
2PL	-áis	-ab-ais	-ate	-av-ate
3PL	-an	-ab-an	-ano	-av-ano

Table 3: Present and imperfect agreement paradigms in Spanish and Italian.

In each language there is one alternation that must be accounted for.

In Spanish, the *-o* in the 1SG present tense slot does not return in the 1SG past tense, where we find *-a* instead. Since *-a* also occurs in the 3SG present tense slot, this form can be analyzed as an elsewhere form. We can then induce impoverishment to account for its occurrence in the 1SG past tense slot with the rule given in (11):

$$(11) [\text{u}\varphi: \text{speaker}] \rightarrow [\text{u}\varphi:] / [\text{T}: \text{past}]$$

This rule deletes the speaker value so that the elsewhere spell-out rule inserting *-a* is used instead of the rule referring to the deleted feature value.

In Italian, we have to account to the fact that *-iamo* in the 1PL present tense context does not return in the past, where we find *-amo* instead. This alternation can be captured with the context-sensitive rule in (12):

$$(12) \text{-amo} \quad \langle \rangle \quad [\text{u}\varphi: \text{speaker}] / [\text{T}: \text{past}]$$

In sum, Spanish and Italian can be straightforwardly analyzed as bi-morphemic languages:

(13) <i>Spanish</i>	<i>Italian</i>
<i>-o</i> $\langle \rangle$ [u φ : speaker]	<i>-o</i> $\langle \rangle$ [u φ : speaker]
<i>-as</i> $\langle \rangle$ [u φ : addressee]	<i>-i</i> $\langle \rangle$ [u φ : addressee]
<i>-a</i> $\langle \rangle$ [u φ :]	<i>-a</i> $\langle \rangle$ [u φ :]
<i>-amos</i> $\langle \rangle$ [u φ : plural, speaker]	<i>-iamo</i> $\langle \rangle$ [u φ : plural, speaker]
<i>-áis</i> $\langle \rangle$ [u φ : addressee, plural]	<i>-ate</i> $\langle \rangle$ [u φ : plural, addressee]
<i>-an</i> $\langle \rangle$ [u φ : plural]	<i>-ano</i> $\langle \rangle$ [u φ : plural]
\emptyset $\langle \rangle$ [T:] -	<i>-amo</i> $\langle \rangle$ [u φ : plural, speaker] / [T: past]
<i>-ab</i> $\langle \rangle$ [T: past]	\emptyset $\langle \rangle$ [T:]
[u φ : speaker] \rightarrow [u φ :] / [T: past]	<i>-av</i> $\langle \rangle$ [T: past]

In contrast to the analyses of the Germanic languages above, the use of an impoverishment or context-sensitive rule does not create additional issues, such as homonymy or null allomorphy, so in that sense there is no temptation to abandon the bi-morphemic analysis of these paradigms.²

3 The relevance of tense and agreement for the licensing of null subjects

In this section, we attempt to relate a bi-morphemic analysis for tense and agreement to successful licensing of null subjects and a mono-morphemic analysis to unsuccessful licensing. This requires two steps. In section 3.1, we will explore what we need to say about parsimony in the analysis of tense and agreement systems so as to make the cut that we need (Germanic languages are mono-morphemic, Romance pro drop languages bi-morphemic). In section 3.2, we will explore why this relationship would exist.

3.1 The choice between a mono-morphemic and bi-morphemic analysis

In section 2, we have seen that the choice between a mono-morphemic and bi-morphemic analysis of tense and agreement forms is not at all straightforward for the Germanic languages whereas it is for the Romance languages. It is not possible, however, to fully derive the choice for each language strictly from the data. Our theoretical toolkit is significant enough to in principle analyze all the languages under scrutiny as bi-morphemic, and one would do so if transparency of tense and agreement forms is an important analytical goal. However, the choice between a mono- or bi-morphemic analysis should not only be theory-, or transparency-driven, but should also be informed by acquisition principles. What is in theory possible may be curtailed by biases brought in by the

² We like to mention that Romanian is also a transparent language in this sense, although the analysis of the paradigm is slightly more complicated. See Koenenman & Zeijlstra (2021) for a detailed analysis.

learner.

Take as an example Icelandic. Although the postulation of homonymous forms allows one to maintain a transparent, i.e. bi-morphemic, analysis of tense and agreement, there is reason to believe that children initially avoid violations of a 1:1 mapping of form and meaning (cf. Mazzocco 1997; Doherty 2004). If so, the postulation of homonyms is not at all innocent from the learner's perspective. Although we do not know of any literature on the acquisition of null allomorphs, it is not unreasonable to think that acquiring these is a significant acquisition burden. After all, there is no independent evidence for the existence of a null allomorph in Standard German, so that its postulation is only informed by the desire to keep the system bi-morphemic.

The choice, then, is the following. If keeping the tense and agreement system morphologically transparent is more important than avoiding homonyms and null allomorphs, the Germanic languages under scrutiny come out as bi-morphemic. If transparency is merely a learner's default assumption, which can be overridden in order to avoid having to postulate homonyms or null allomorphs, these languages will be mono-morphemic. Under either position, Spanish and Italian should come out as bi-morphemic and, since impoverishment and context-sensitive rules are required to obtain that result, these devices cannot be decisive. Now, in the absence of clear acquisition evidence one way or the other, the task is to see what each position buys us. As section 3.2 points out, if we take transparency not to be a theoretical restriction but merely an initial, default assumption, it enables us to understand why Icelandic, Standard German and English do not have argumental null subjects.

3.2 The overspecification problem

Naturally, the question comes up why the distinction between mono- and bimorphemic tense and agreement markers should correlate with the presence or absence of classical pro drop. We argue that a mono-morphemic analysis leads to a problem of overspecification.

There is a general consensus in the literature that agreement underspecification is a problem for the licensing of null subjects. In a contextual approach to pro drop, this means that in a particular context, e.g. first person plural, a morpheme with underspecified agreement information will not successfully lead to the reconstruction of the unexpressed subject. We propose is that overspecification gives rise to a similar effect. A morpheme that does not only contain agreement but also tense features cannot successfully license an empty subject. This means that there are four scenarios to consider. These are given in Table 4.

Scenario	Subject	Morpheme (on V)	Status	Pro drop
1	[plural, speaker]	[plural, speaker]	specification	yes
2	[plural, speaker]	[plural]	underspecification	no
3	[plural, speaker]	[plural, speaker, past]	overspecification	no
4	[plural, speaker]	[plural, past]	under- & overspecification	no

Table 4: Under- and overspecification scenarios and the licensing of null subjects.

Scenario 1 reflects a context that licenses a 1PL null subject, as the information on the subject and licensing morpheme are the same. Scenario 2 constitutes a case of underspecification: the licensing morpheme lacks one feature and this bleeds pro drop. Under our proposal, null subjects are also impossible in scenario 3. Here, the licensing morpheme shares all its features with the subject but it crucially also expresses a tense feature. Scenario 4, finally, reflects a theoretical possibility that does not arise in our limited data set.

The question then is why overspecification of an agreement morpheme should result in the absence of pro drop. The answer to this question depends on the theoretical perspective on pro drop one may have in general. There are essentially two theoretical approaches, both defended in the literature: either the agreement stands in some kind of feature-sharing relation with an unpronounced pronominal subject, dubbed *pro*, or, even stronger, the agreement marker *is* the subject pronoun.

The first position postulates the existence of an empty subject pronoun that syntactically takes the same position as an overt one. The difference between languages like Italian and English

is then minimal: Both have subjects agreeing with the verb but in Italian this subject can be covert, *pro*. As there is only one *pro*, it must be featurally underspecified in the lexicon. At the same time, *pro* must end up carrying the relevant ϕ -features, as otherwise it cannot be interpreted as a real pronoun at LF. The way to make *pro* receive those ϕ -features, then, is by having the agreement markers themselves value it. However, the above does not explain yet why in languages like Icelandic and Standard German, in which the mono-morphemic expression of tense and agreement entails that these properties are hosted on the same projecting functional head, the agreement marker cannot license pro drop. After all, why couldn't the ϕ -values on a functional head I° simply value the ϕ -feature of the subject DP.

In a language like Icelandic, however, tense and agreement are subfeatures of a feature [I] and this has non-trivial consequences for the content of the agreement subfeature. Since tense is an interpretable subfeature, I itself must be interpretable. But this means that the agreement subfeatures must be interpretable too. The single I-feature [_I T: past; ϕ : speaker, plural] cannot appear fully valued in the derivation prior to the split to LF. If it were, both the ϕ - and the tense subfeature would feed interpretation at LF, which would make the sentence crash: tense and ϕ -features cannot be interpreted in the same semantic position, because they are of different pragma-semantic types. The only way for the relevant I-feature to circumvent LF-crashing is by being lexically unvalued for either tense or ϕ , not for both. Since tense feature values are not present anywhere else in the derivation, those values must be directly inserted from the lexicon; otherwise, a past tense value would not be visible at LF. But if tense must be valued on the relevant I-subfeature, ϕ cannot be. Hence, I(NFL)-features that comprise both tense and ϕ -subfeatures must have their ϕ -subfeatures valued in the course of derivation. To come back to the Icelandic example, only [_I T: past; ϕ : _ _] can be a lexical item, not [_I T: past; ϕ : speaker, plural]. But without values for the agreement subfeature, this I-feature is unable to license *pro*. This explains why under approaches licensing a covert *pro*, overspecification is not allowed.

Our approach is also compatible with the alternative approach, in which rich agreement markers are taken to be actual pronouns, i.e. weak subject markers, whose rich ϕ -features enable them to be interpreted as such. This is the position adopted by Barbosa (2009) (see also Borer 1986, and Alexiadou & Anagnostopoulou 1995, among others, for similar proposals). If agreement markers are weak pronouns they should have exactly those features that overt (weak) pronouns exhibit as well. As such pronouns never exhibit any tense morphology, it is predicted that such subject agreement markers cannot be featurally overspecified either.

4 Partial pro drop

One immediate question that comes up is how to deal with partial pro drop languages. We have shown that it is possible in a contextual approach to derive paradigmatic effects. If two forms, A and B, compete for insertion and target the same morpheme, A and B must express the same type of feature(s). If B expresses tense *and* agreement, so must A and by transitivity all other forms that A competes with. The consequence of this is that the entire paradigm is analyzed in mono-morphemic terms and pro drop is effectively blocked across the board. This gives the right result for Icelandic, Standard German and English but begs the question of how to deal with languages that allow pro drop in certain but not all contexts. We have nothing to say at the moment about famous cases like Standard Finnish and Hebrew but think that the comparison between Standard German and Germanic dialects proves insightful.

In a contextual approach to pro drop it is in principle possible that null subject licensing is not uniform across the whole paradigm but restricted to certain contexts. Given the logic about transitivity, there must be something that stops the spread of the mono-morphemic analysis over all slots of the paradigm. Let us take a look at Frisian. Like Standard German, the 3SG form does not return in the past tense (see Table 5). Frisian must be mono-morphemic and pro drop is not predicted. This expectation is correct for all slots of the paradigm except for the 2SG slot, where pro drop is possible (cf. 14, from De Haan 1984). It is noteworthy that it is exactly in 2SG contexts that a second agreement marker pops up, namely on the complementizer, as illustrated in (14b).

	Frisian	
	Present	Past
1SG	-∅	-te
2SG	-st	-te-st
3SG	-t	-te/*tet
1PL	-e	-te-(e)n
2PL	-e	-te-(e)n
3PL	-e	-te-(e)n

Table 5: Frisian tense and agreement paradigm.

- (14) a. Ik denk dat-st (do) my helpe moatst
 I think that.2SG me help must
 ‘I think that you should help me.’
 b. Miskien moat-st (do) my helpe
 Perhaps must.2SG (you) me help
 ‘Perhaps you should help me.’

The correlation between partial pro drop and complementizer agreement is a known one. Rosenkvist (2009: 163) for instance notes that in Bavarian dialects it is exceptionlessly the case that pro drop only takes place in contexts that have complementizer agreement, which is often the 2SG context but sometimes also the 1PL or 2PL context (see also Fuss 2005).³ From our perspective, this raises two questions: (i) why would complementizer agreement lead to a bi-morphemic analysis for tense and agreement for the context it appears in, and (ii) why would null subjects also be licensed for the same contexts in main clauses, where no complementizer is present (cf. 14b)?

The answer to the first question is that complementizer agreement provides clear evidence for a bi-morphemic analysis of tense and agreement for the context it appears in. There are two noteworthy properties of the complementizer in Frisian and Bavarian. First of all, the form of the agreement marker is identical to the one appearing on the verb. Second, the complementizer is never marked for tense. This means that there must be an underlying morpheme with the sole purpose of expressing agreement. Given the phonetic resemblance of the agreement marker on the verb and complementizer, the following spell-out rule would generalize over both instances:

- (15) *-st* <> [u∅: addressee]

The consequence of this analysis would be that in this 2SG context the tense features expressed on the verb must be part of a separate morpheme dedicated to tense. In this way, the overspecification problem is now circumvented, as the morpheme expressing [u∅: addressee] can straightforwardly license a 2SG null subject. It is also clear what the answer to question (ii) should be. It is not the complementizer itself that licenses the null subject in (14a). Complementizer agreement provides evidence for the bi-morphemic analysis of tense and agreement in 2SG contexts, and for the existence of a separate agreement morpheme. It is this morpheme that licenses the null subject, both in embedded clauses (cf. 14a) and main clauses (14b). Finally, it is also clear why Standard German has no partial pro drop. It lacks complementizer agreement, which constitutes the crucial evidence for a bi-morphemic system.

³ Bavarian dialects raise an additional question. Since they generally lack a past tense paradigm but use a compound tense instead, they do not display the cue for a mono-morphemic analysis, namely the non-appearance of the 3SG present tense form in the past. This raises the question why these dialects are not bi-morphemic and allow null subjects in at least 1SG contexts, which display a unique agreement form on the verb. See Koeneman & Zeijlstra (2021) for an argument that the absence of a past tense entail mono-morphemicity.

5 Conclusions

Despite the enormous attention that pro drop has received in the linguistic literature, there is no generally accepted answer to the question why relatively rich Germanic languages do not have argumental null subjects, neither is there a fundamental answer to the question why English would not allow them in at least 3SG contexts. This may suggest that at least one of our working hypotheses is misguided. The purpose of this paper is to show that there is something to gain from abandoning the hypothesis that bi-morphemicity is a goal in the analysis of tense and agreement paradigms, namely a way of relating the mono-/bi-morphemic distinction to the presence or absence of null subjects.

We argued that featural richness (i.e., agreement markers that allow reconstruction of the subject) is a necessary but not a sufficient condition for licensing pro drop. Pro drop is only possible if tense and agreement markers are bi-morphemic. In an interplay of grammatical considerations and learning principles, Icelandic and Standard German are likely to be languages in which only one morpheme carries both tense and ϕ -features. We also argued that existing approaches to the nature of pro drop actually require of agreement markers that they are not featurally overspecified, thus embedding our conclusions in a larger theoretical picture. An empirical prediction that follows from our analysis is that every language in which the elsewhere agreement marker does not reappear in any of the other paradigms cannot be a classical pro drop language.

Finally, we argued that, even though single paradigms ought to be generally uniform when it comes to mono-/bimorphemicity, exceptions can be made in particular cases. This is what underlies the nature of partial pro drop in Germanic dialects.

References

- Alexiadou, Artemis & Elena Anagnostopoulou. 1998. Parametrizing Agr: Word Order, V-Movement, and EPP-Checking. *Natural Language and Linguistic Theory* 16: 491–539.
- Barbosa, Pilar. 1995. Null Subjects. Doctoral dissertation at Massachusetts Institute of Technology at Cambridge.
- Benveniste, Emile. 1971. *Problems in general linguistics*. Coral Gables, FL: University of Miami Press.
- Bobaljik, Jonathan. 2003. Realizing Germanic inflection: why morphology does not drive syntax. *Journal of Comparative Germanic Linguistics* 6: 129–167.
- Borer, Hagit. 1986. I-Subjects. *Linguistic Inquiry* 17, 375–416.
- de Haan, Ger. 1994. Inflection and cliticization in Frisian. *NOWELE* 23: 75–90.
- Doherty, Martin J. 2004. Children's difficulty in learning homonyms. *Journal of Child Language* 31(1): 2003–214.
- Fuß, Eric. 2005. *The Rise of Agreement*. Amsterdam/Philadelphia: John Benjamins.
- Garbacz, Piotr. 2010. *Word order in Övdalian: a study in variation and change*. Centre for Languages and Literature, Lund University.
- Harley, Heidi and Elizabeth Ritter. 2002. A feature-geometric analysis of person and number. *Language* 78: 482–526.
- Koenenman, Olaf. 2000. The flexible nature of verb movement. Doctoral dissertation, Utrecht University.
- Koenenman, Olaf and Hedde Zeijlstra. 2021. Deriving pro drop in a non-paradigmatic way. Ms., Radboud University at Nijmegen and Goettingen University.
- Mazzocco, Michele M. 1997. Children's interpretations of homonyms: a developmental study. *Journal of Child Language* 24(2): 441–467.
- Neeleman, Ad and Kriszta Szendrői. 2007. Radical pro drop and the morphology of pronouns. *Linguistic Inquiry* 38: 671–714.
- Preminger, Omer. 2014. *Agreement and its failures*. Cambridge, MA: The MIT Press.
- Rosenkvist, Henrik. 2009. Referential Null Subjects in Germanic Languages – an Overview. *Working Papers in Scandinavian Syntax* 84: 151–180.
- Tamburelli, Marco. 2006. Remarks on richness. *UCL Working Papers in Linguistics* 18: 439–455.

Radboud University Nijmegen
English Language and Culture
Erasmusplein 1
6525 HT Nijmegen
The Netherlands
olaf.koeneman@ru.nl

University Goettingen
Seminar for English Philology English language and culture
Käte-Hamburger-Weg 3
37073 Göttingen
Germany
hzeijls@uni-goettingen.de