Pro drop and the morphological structure of inflection

Olaf Koeneman & Hedde Zeijlstra

Abstract

A central problem for pro drop theory is how to account for the distribution of argumental null subjects. A paradigmatic approach, in which the whole paradigm determines whether a language can have null subjects or not, undergenerates in that it does not predict the existence of partial pro drop languages. An approach in which the licensing of null subjects is determined per context overgenerates in that it is not obvious why e.g. English does not have null subjects in at least the 3rd person singular. In this article, we aim at repairing the overgeneration problem for contextual approaches. By contrasting the Romance pro drop languages with the Germanic languages and Standard French, we argue that only the latter express tense and agreement with the same underlying morpheme, which therefore contains features (tense) that are incompatible with the subject it would have to license: it is overspecified. The question is then what determines the choice between a bi- or mono-morphemic expression of tense and agreement, and we will argue that reasonable assumptions about the acquisition of morphological systems makes the right cut.

1 Introduction

Some languages allow the argumental subject of a finite clause to be (phonologically) absent, whereas this leads to ungrammaticality in other languages. This difference can be illustrated with Italian and English.

(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*

Italian is a so-called pro drop, or null subject, language, in contrast to English. The classical question is what underlies this distinction, and the classical answer is that the source lies in the

inflectional agreement systems (cf. Taraldsen 1978, Rizzi 1982, Jaeggli & Safir 1989 for the original formulations of this conjecture). Focusing on Italian and English, one can observe that English is a poor agreement language, only displaying a distinction between the 3SG *-s* and a null form (in the present tense), whereas in Italian each slot of the paradigm is filled by a unique form:

(2) Italian and English agreement paradigms:

	Italian	*English*
	parlare ('to talk')	to talk
1SG	parl-o	talk-∅
2SG	parl-i	talk-∅
3SG	parl-a	talk-s
1PL	parl-iamo	talk-∅
2PL	parl-ate	talk-∅
3PL	parl-ano	talk-∅

The intuition behind the correlation is straightforward. With a rich agreement system, the unexpressed subject can be straightforwardly identified. However, if identification is what is at stake, reference to richness is not enough. After all, the 3SG form *-s* is as unique to the English paradigm as the *-a* is to the Italian one. The question is therefore why English does not at least allow null subjects in 3SG contexts. A common strategy is to refer to the entire paradigm: the Italian paradigm is rich overall and therefore allows null subjects across the board, whereas the English paradigm is poor, thus blocking null subjects across the board (Rizzi 1982, Jaeggli & Safir 1989). Reference to the paradigm may even help to account for the fact that pro drop is absent in rich agreement languages like Icelandic and Standard German. As can be observed in (3), both the Icelandic and Standard German paradigms display syncretisms (the bold-faced forms), which

could be a factor that renders these paradigms poor for the purposes of null subject licensing (cf. Koeneman 2000; Tamburelli 2006).

(3) Icelandic and Standard German (present tense) agreement paradigms:

	Icelandic	*Standard German*
	heyra ('to hear')	spazieren ('to walk')
1SG	heyr-i	spazier-e
2SG	**heyr-ir**	spazier-st
3SG	**heyr-ir**	**spazier-t**
1PL	heyr-jum	**spazier-en**
2PL	heyr-ið	**spazier-t**
3PL	heyr-a	**spazier-en**

There are several problems with the paradigmatic approach, however.

First of all, Romanian is an example of a language that also displays syncretism in its paradigm (in the first conjugation in the 3SG and 3PL cells, cf. (4)). A look at the past tense paradigms of Romanian and Icelandic reveals a further similarity: what is syncretic in the present, is not syncretic in the past, and vice versa.

(4) Romanian and Icelandic paradigms:

	Romanian		*Icelandic*	
	cânta ('to walk')		heyra ('to hear')	
	present	*past (imperfect)*	*Present*	*Past*
1SG	cânt	cântam	heyr-i	heyrði
2SG	cânţi	cântai	**heyr-ir**	heyrðir
3SG	**cântă**	cânta	**heyr-ir**	heyrði
1PL	cântăm	cântam	heyr-jum	heyrðum

2PL	cântaţi	cântaţi	heyr-ið	heyrðuð
3PL	**cântă**	cântau	heyr-a	heyrðu

This means that - overall - contrasts are made between 1st, 2nd and 3rd person singular and plural forms in both languages. This is apparently not sufficient because, despite these similarities, only Romanian is a full-fledged null subject language.

A second problem for the paradigmatic approach is the existence of so-called partial pro drop languages, where pro drop is possible in some but not all slots of the paradigm. These strongly suggest that reference to the whole paradigm is unnecessary for pro drop to be possible. Standard Finnish and Hebrew are famous examples but closer to the languages put central in this paper, the Germanic ones, we can observe that Frisian and Bavarian German allow pro drop in 2SG contexts, and some Bavarian varieties also in 1PL or 2PL contexts. These varieties, then, defy an all-or-nothing (i.e. parametric) approach to pro drop.

Third, from a theoretical perspective, paradigms are epiphenomenal in the sense that they are handy devices for linguists or someone learning the language but they are not constructs that a grammatical rule should be able to refer to (cf. Bobaljik 2003). If so, a paradigmatic theory of pro drop is hard to even formulate.

Given these arguments against a paradigmatic approach to pro drop, one could revert to a contextual approach, in which the possibility of having a null subject must be determined for each context individually. That would potentially provide a handle on partial pro drop languages but beg the question why Icelandic and Standard German, or even English, would not at least have partial pro drop in contexts with a unique agreement form. In short, if we compare the paradigmatic approach with the contextual approach, we can state that the former undergenerates (in that it does not predict partial pro drop), whereas the latter overgenerates (in that it predicts more partial pro drop than we actually find). This means that the following empirical questions stand:

(i) Why do languages like Italian, Spanish and Romanian, but not Icelandic and Standard German, allow full pro drop?

(ii) Why do languages like English, Icelandic and Standard German not at least allow partial pro drop?

(iii) Why do languages like Frisian and Bavarian dialects allow partial pro drop?[1]

Given the discussion above, we conclude that featural richness is a necessary but not sufficient prerequisite for the licensing of pro drop. What, then, is the additional factor? We propose that not only should an empty subject be licensed by agreement features that are rich enough to allow reconstruction of the unexpressed subject, it should also be licensed by a morpheme that *only* expresses agreement features. If the same morpheme also expresses tense features, this morpheme is overspecified for the job. The reason why Italian *-a* licenses pro drop but English *-s* does not is then that *-a* only realizes agreement features while *-s* realizes both tense and agreement features. Before we can arrive at this conclusion, however, we must look in detail into the relevant morphological paradigms. This is taken on in section 3, after we have fleshed out our proposal in a bit more detail in section 2. Our conclusion will be that a bi-morphemic analysis of tense and agreement is straightforward in Spanish and Italian but that the choice between a bi- or mono-morphemic analysis is harder to make for the Germanic languages and Standard French. In section 4, we will explore what we need to assume in order to maintain a correlation between bi-morphemicity (pro drop) and mono-morphemicity (no pro drop) and show that this can be achieved by making reasonable assumptions about how the child acquires morphological systems. Section 5 deals with the overspecification problem in more theoretical depth (5.1) and deals with the problem of partial pro drop languages (5.2). Section 6 concludes.

[1] In order to keep the typological distance minimal, we will only look at German dialects and Frisian in section 5.2, and refrain from an analysis of Standard Finnish and Hebrew. The latter two constitute a different kind of partial pro drop language, one that disallows argumental pro drop in 3SG contexts except in embedded clauses (cf. Vainikka & Levy 1999; Holmberg 2005; see also Section 6).

2 Overspecification as a problem for pro drop licensing

An uncontroversial pro-drop language like Italian can be characterized as a rich agreement language. For every person/number context, it has at least one pronoun and each of these pronouns corresponds with a distinct agreement form. It has enough contrasts to allow the postulation of agreement features like [speaker], [addressee] and [plural]. It is also a variety in which tense and agreement are expressed by distinct morphemes. Take the verb form *amavo* ('I loved'), in which *-(a)v-* is the imperfect marker and *-o* the 1SG marker. A language without pro drop like English can be uncontroversially characterized as a poor agreement language. It only has *-s* and a zero form and therefore does not generate a [speaker] or [addressee] feature. It can also be characterized by the fact that tense and agreement are not expressed bi-morphemically. A strong indication for this is the fact that tense and agreement forms never co-occur overtly: *John walks*, *John walked*, **John walkeds*. Bobaljik & Thráinsson (1998) argue on the basis of this that English has a so-called unsplit IP. There is only one functional head in the clausal spine, INFL, responsible for the expression of tense and agreement. As a consequence, *-s* and *-ed* compete for the same slot. Italian, then, can be analyzed as a split-IP language, with dedicated slots for the expression of tense and agreement.

Given this distinction, it becomes tempting to relate the split-IP parameter to the pro drop distinction: For pro drop to be possible, the syntax must generate a dedicated agreement head. Since English only has INFL, even the *-s* form is incapable of licensing a null subject, despite being a unique marker for the 3SG context. But why would the split-IP parameter be relevant? Under the assumption that richness somehow correlates with pro drop licensing, four scenarios have to be evaluated. Consider the hypothetical case of a first person plural subject, expressing the features [plural, speaker]:

(5) Hypothetical scenarios:

Context	*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- and overspecification	No

In context 1, the features expressed by the morpheme on the verb coincide with those expressed by the subject. This means that an empty subject should be possible because in context 1 the morpheme on the verb expresses the same features as those of the subject (and no more) and is therefore rich enough to allow subject identification. Context 2 involves underspecification, because the licensing morpheme fails to express one feature that the subject expresses. We therefore predict pro drop not to be possible here. Context 3 involves overspecification: the morpheme on the verb expresses more features than the subject does. Using an empty subject in this context is not possible, we propose, as the overspecific feature [past] is crucially one that cannot be part of the interpretation of the subject.[2] As a consequence, no empty subject can be used. It is now logically possible to have a context with both under- and overspecification (one feature too many and one feature too few), and no null subjects should be possible in Context 4.[3]

The overspecification problem arising in Context 3 disappears if the language in question expresses tense on a separate morpheme, as illustrated in (6). below. In such cases, there is one morpheme, namely Morpheme 1, that can adequately perform the licensing function because it is

[2] The question arises why the tense features cannot simply be ignored. We will discuss this issue in section 5.1.

[3] For some poorly understood reason, the licensing morpheme can be underspecified for gender. In 3rd person contexts, where gender marking occurs, pro drop is possible in a language like Italian, despite the fact that this language lacks gender marking on the finite verb. We speculate that the reason for this is that gender is not a general property of subjects, and its irrelevance in 1st and 2nd person contexts carries over to 3rd person contexts. Note that overspecification of gender should not be a problem, as gender can be interpreted as part of the subject, in contrast to tense. Therefore, pro-drop languages with gender marking only on the verb would be compatible with our analysis.

neither under- nor overspecified. The feature [past] does not interfere with null subject licensing because it is expressed by a separate morpheme.

(6) A bi-morphemic scenario:

Subject	*Morpheme 1*	*Morpheme 2*	*Status*	*Prodrop*
[plural, speaker]	[plural, speaker]	[past]	Specification	Yes

Given this proposal, the next question is how to determine exactly for each language under discussion whether it expresses tense and agreement bi- or mono-morphemically. This task will be undertaken in the next section.

3 Bi-morphemic versus mono-morphemic expression of tense and agreement

Let us see what evidence languages provide for separate agreement morphemes. We will start with polar opposites, the rich agreement languages Italian and Spanish (section 3.1) versus the poor agreement languages English, Dutch and Faroese (section 3.2), where the choice between a split and an unsplit IP (and thus for separate morphemes or not) is always based on regular verbal paradigms, in line Bobaljik & Thrainsson (1998); this essentially means that for languages like English we will discard modal paradigms or the paradigms of *be/have*. After this comparison, we are in a better position to determine the status of rich agreement languages without pro drop, Icelandic, Standard German and Standard French (section 3.3).

3.1 Rich agreement languages with pro drop: Spanish and Italian

Spanish and Italian are textbook pro drop languages, which is usually tied to the fact that they are rich agreement languages. Rich agreement is not a sufficient prerequisite for pro drop, though, so let us see in what way they express tense and agreement in a transparent way. We must look at non-present tenses to establish this, since present tense is usually not expressed overtly in these languages. Take, therefore, the paradigms of the imperfect tense:

(7) *Present and imperfect agreement paradigms in Spanish and Italian*

	Spanish		Italian	
	amar ('to love')		amare ('to love')	
	present	imperfect	present	imperfect
1SG	amo	amaba	amo	amavo
2SG	amas	amabas	ami	amavi
3SG	ama	amaba	ama	amava
1PL	amamos	amábamos	amiamo	amavamo
2PL	amáis	amabais	amate	amavate
3PL	aman	amaban	amano	amavano

Both languages have a consonant marking the imperfect, *-b-* and *-v-* respectively, and these are followed by agreement endings that by and large also appear in the present tense, rendering a bi-morphemic analysis straightforward.

In our analyses of these paradigms, we assume that only privative features can be postulated (cf. Bonet 1991, Harley & Ritter 2002, Béjar 2003 and Nevins 2003 for discussion) and that unmarked features reflect the absence of a particular feature value (see Sauerland 2002, 2008, Heim 2008, Kratzer 2009). This entails that singular constitutes the absence of a number value and present tense the absence of a tense value. In the same vein, the 3rd person is generally taken to be the non-person defined by the absence of person values (cf. Forschheimer 1953; Benveniste 1971; Harley & Ritter 2002; Preminger 2014). This entails that the form appearing in the 3SG slot counts as the elsewhere form, a fairly standard and uncontroversial assumption, but one that will play a significant role in our analysis.

We further follow Oltra-Massuet (1999) in assuming that the Spanish-*a-* vowels following the stem are theme vowels, as they differ per conjugation.[4] Oltra-Massuet argues that the same is

[4] We will refrain from a complete analysis of theme vowels, as they do not influence the decision between a split- or unsplit analysis of tense and agreement. A full understanding of theme vowels must include an analysis of their (non)-

true for the *-a-* vowel following the *-b-* in the Spanish imperfect. Under this analysis, the 3SG present tense slot is always marked with a null form. These assumptions also carry over to Italian (cf. Calabrese 2015, 2019 for details), which leads to the bimorphemic analyses in (8) and (9):

(8) Spanish

Agreement:			*Tense:*		
-o	<>	[uφ: speaker]	-Ø	<>	[T:]
-s	<>	[uφ: addressee]	*-b-*	<>	[T: past]
-Ø	<>	[uφ:]			
-mos	<>	[uφ: speaker, plural]	*Theme vowel:*		
-is	<>	[uφ: addressee, plural]	*-a*	<>	[TH] / [Conjugation 1]
-n	<>	[uφ: plural]			

(9) Italian

Agreement:			*Tense:*		
-o	<>	[uφ: speaker]	-Ø	<>	[T:] (present)
-i	<>	[uφ: addressee]	*-v-*	<>	[T: past]
-Ø	<>	[uφ:]			
-iamo	<>	[uφ: speaker, plural]	*Theme vowel:*		
-te	<>	[uφ: addressee, plural]	*-a*	<>	[TH] / [Conjugation 1]
-no	<>	[uφ: plural]			

As can be established, there is no form that spells out a mixture of tense and agreement features. There are clear tense markers and clear agreement markers: the tense markers only covary with grammatical tense, and the agreement markers only covary with person and number.

occurrence (note e.g. that in the present tense the theme vowel is deleted before a [-high] vowel, like *-o*, in order to block 1SG *amáo* but allow 2PL *amáis*) and of how they change across paradigms. See Arregi (2000) for a concrete proposal for Spanish.

Before continuing, we need to consider two cases in more detail, however. In Spanish, the agreement marker in the 1SG present tense slot, *-o*, does not reappear in the imperfect, where we see *-a* instead. For Italian, we observe that the *-i-* we see in the 1PL present disappears in the imperfect.

What happens in the Spanish imperfect is that the 1SG *-o* is replaced by the elsewhere null form. This can be achieved by inducing the following impoverishment rule:

(10) [uφ: speaker] > [uφ:] / [T: past]

This rule deletes the [speaker] value in the past tense so that *-o* can no longer be inserted. The elsewhere agreement form -∅ is inserted instead. This gives a 1SG-3SG syncretism that we also find elsewhere in Spanish conjugation (e.g. in the indictive conditional and the subjunctive), and ensures that the formal similarity of 1SG and 3SG contexts is not coincidental but structural.

Let us now turn to disappearance of the 1PL *-i-* in the Italian 1PL imperfect context. One analysis is that the agreement endings are *-iamo* and *-amo* respectively. Insertion of *-amo* in the past tense can then be ensured by adopting the context-sensitive rule in (11):

(11) *-amo* <> [uφ: speaker] / [T: past]

Note that this does not entail that *-amo* spells out a past tense feature, but only that [T: past] defines the context of application for this rule.

Calabrese (2019) analyzes the *-iamo*/*-amo* alternation as allomorphy involving theme vowels. The affix *-mo* is analyzed as the 1PL agreement marker preceded by the theme vowel *-a-* in imperfect 1PL contexts but by *-ya-* in the present tense. Under such an analysis, the point remains the same: The alternation can be captured by a context-sensitive rule and no switch to a mono-morphemic analysis is required. Hence, impoverishment rules and context-sensitive allomorphs naturally uphold the bi-morphemic analysis for Spanish and Italian. [5]

[5] A look at other conjugations reveals that Italian is more transparent than Spanish. Whereas the Italian *-v-* imperfect marker returns in other conjugations (the *-ere* and *-ire* classes), the Spanish *-b-* does not. In the *-er(e)* conjugation for instance, Italian has forms like *credevo* (believe.IMPERF.1SG), whereas Spanish has *termía* (fear.IMPERF.SG). Under

It must be acknowledged that Spanish and Italian are not transparent all the way down, though. They also have synthetic perfectives, known as the *pretérito perfecto* and *passato remoto*, respectively. The forms for the first conjugation are given in (12).

(12) *Spanish pretérito perfecto* and *and Italian passato remoto in the first conjugation*

	Spanish	Italian
1SG	amé	amai
2SG	amaste	amasti
3SG	amó	amò
1PL	amamos	amammo
2PL	amasteis	amaste
3PL	amaron	amarono

There is no consonant marking this tense in either language, as in the imperfect, and the vowel following the stem is similar to the one we see in the present tense and could therefore be analyzed as a theme vowel. To account for the perfective semantics, we must either assume that it is expressed by a null form or, as Arregi (2000) proposes for Spanish and Calabrese (2019) for Italian, by a fused tense and agreement form (cf. for fusion Halle & Marantz 1993: 136). Under the latter analysis, the implication is still that tense and agreement are encoded as separate nodes in the syntax but that the two nodes fuse post-syntactically to allow insertion of a single morpho-phonological form expressing both tense and agreement features. Although we are not forced to

the assumption that the vowels following the stem are theme vowels, the imperfect marker is overt in the *-ar* conjugation but null elsewhere: term-i-Ø-a (cf. Arregi 2000). The null form has a wider distribution in that it occurs in all but the first conjugation, and within the first conjugation also shows up in the past tense of the future, where we find *amaria* (love.FUT.IMPERF.SG) instead of the expected *amaraba.* One might therefore treat the null form as the unmarked (or elsewhere) imperfect marker and the *-b-* as one whose use must be restricted to the imperfect of the non-future. Note, however, that an analysis that postulates a null imperfect tense marker leaves it as an accident that the tense marker is null whenever it follows high theme vowel *-i*. Another approach would therefore be to assume that the imperfect marker is *-b-* throughout but that it gets phonologically deleted after *-i*. We will not pursue this issue further.

commit ourselves to either analysis, a question that arises under the fusion analysis is why the existence of these synthetic perfectives would not turn Spanish and Italian into mono-morphemic languages. The answer must lie in the robust evidence against mono-morphemicity elsewhere.[6] If some paradigms provide robust evidence for separate tense and agreement morphemes and other paradigms can be analyzed using fusion, there is no reason to abandon a bi-morphemic analysis underlying all paradigms.

To sum up, Spanish and Italian are highly amenable to a bi-morphemic analysis of tense and agreement. This does not mean, however, that a mono-morphemic analysis is impossible. If one so wishes, one could analyze Spanish as having an unsplit-IP, with a form like *-b(a)s* spelling out a mixture of tense and agreement features. This would of course be at the expense of losing morphological transparency. It is only under a bi-morphemic analysis that *-s* can be listed as a separate vocabulary entry, inserted in 2SG contexts in present and imperfect tenses alike. as If null subject licensing correlates with bi-morphemicity, it must be the case that a bi-morphemic analysis is the default one and that it requires a certain trigger to abandon it in favor of a mono-morphemic analysis. The next section will identify those triggers.

3.2 Poor agreement languages without pro drop: English, Dutch and Faroese

In the previous section, we have looked at two rich agreement languages. In contrast, English is a poor agreement language in that it does not display enough morphological contrasts for the postulation of [speaker], [addressee] and [plural]. The same is true for Dutch and Faroese, which lack a contrast between 1st and 2nd person (see (16)).

Recall that under a contextual approach to pro drop having a poor agreement paradigm does not entail the absence of pro drop altogether (recall the discussion about the English *-s*).

[6] For Italian, the answer is also that the *passato remoto* is mostly obsolete and not part of the register naturally acquired by speakers of the standard language, but this is not the case for Spanish and certain Italian dialects such as Sicilian or Tuscan (cf. Ledgeway 2016).

Under the approach pursued here, it therefore needs to be independently established if these languages are bi- or mono-morphemic. As already discussed earlier, English provides one clear indication that it is mono-morphemic, namely the fact that its only overt agreement ending (3SG *-s*) never co-occurs with the past tense ending. Bobaljik & Thráinsson (1998) take this to mean that the two forms are in complementary distribution, which then reveals the existence of one morphological slot in which either *-s* or *-ed* must be inserted. Straightforward as the analysis is, let us nevertheless establish what a bi-morphemic analysis of English would look like. Under the assumption that the 3SG form is the elsewhere, the following spell-out rules could then be postulated as a starter:

(13) English

Agreement:			*Tense:*		
-Ø	<>	[uφ: participant]	-Ø	<>	[T:]
-s	<>	[uφ:]	*-ed*	<>	[T: past]
-Ø	<>	[uφ: plural]			

As it stands, however, (13) wrongly generates the *-s* in the past tense. In order to block the *-s* from appearing in the past tense we cannot make use of an impoverishment rule. As the 3rd person form is the elsewhere, there are simply no feature values to impoverish. This means that the only solution is to enrich the grammar with yet another null form, an allomorph that is contextually restricted to occur in past tense contexts only:

(14) -Ø <> [uφ:] / [T: past]

Now, let's compare this analysis with the mono-morphemic one in (15):

(15) English

Inflection:

-Ø	<>	[T:], [uφ: participant]
-s	<>	[T:], [uφ:]

-Ø <> [T:], [uφ: plural]

-ed <> [T: past]

In order to uphold a bi-morphemic analysis for English, then, one must double the number of null forms and ignore what looks like a straightforward complementary distribution effect of two overt forms, *-s* and *-ed*. In addition, one must postulate the existence of four (instead of two) homonyms, forms that look the same (namely null) but spell out different features. From the perspective of the number of null forms, the number of homonymous forms and number of rules, the mono-morphemic analysis is thus by far the most parsimonious one, underscoring Bobaljik & Thráinsson's analysis, although it remains to be determined how these factors weight in exactly, a task we will carry out in Section 5.

Turning now to Dutch and Faroese, we observe that they share one crucial property with English, namely that the form occurring in the 3SG slot does not reappear in the 3SG past tense slot. In Dutch, *-t* is dropped and the form we see in the 3SG past tense slot occurs in every singular slot. In Faroese, *-ir* does not reappear but we see *-i* instead.

(16) English, Dutch and Faroese present and past tense paradigms:

	English		*Dutch*		*Faroese*	
	to talk		danken ('to thank')		døma ('to judge')	
	present	past	present	past	present	past
1SG	talk-Ø	talk-ed	dank-Ø	dank-te	døm-i	døm-di
2SG	talk-Ø	talk-ed	dank-t	dank-te	døm-ir	døm-di
3SG	talk-s	talk-ed/*-eds	dank-t	dank-te/*-tet	døm-ir	døm-di/*-dir
1PL	talk-Ø	talk-ed	dank-en	dank-ten	døm-a	døm-du
2PL	talk-Ø	talk-ed	dank-en	dank-ten	døm-a	døm-du
3PL	talk-Ø	talk-ed	dank-en	dank-ten	døm-a	døm-du

In order to account for the absence of the 3SG agreement form in the past tense, impoverishment is unavailable, as there is no feature value to impoverish. The two alternatives that can be pursued are (i) postulating an elsewhere allomorph for the 3SG past tense context or (ii) assuming that the overt agreement form in the 3SG present tense competes directly with the past tense form in the 3SG past tense, i.e settle for a mono-morphemic analysis. Let us look at these options in detail.

Option (i) allows us to maintain a bi-morphemic analysis, but for Dutch comes at the cost of having to postulate a contextual null allomorph. In addition, the form that consequently fills the 3SG slot in the past must be accidentally homonymous with the 1SG past form, because it shares with the 1SG agreement form that it is null. In Danish, the contextual allomorph would be overt (namely *-i*) but like in Dutch would be accidentally homonymous to the *-i* occurring in 1SG past tense contexts. These accidental homonymies can be circumvented by impoverishing the speaker feature in the past, so that the same contextual allomorph is inserted in all past singular contexts, but this would amount to an overall complication of the grammar in order to achieve a simplification.

Option (ii) boils down to letting the agreement forms in the present tense compete directly with past tense forms. Like *-s* competes with *-ed* in English, Dutch *-t* competes with *-de* and Faroese *-ir* competes with *-di*. The explanation for why *-s*, *-t* and *-ir* do not reappear in the past tense is then straightforward: competing forms cannot be stacked. This analysis requires that these forms target the same morpheme for insertion and entails a mono-morphemic analysis, in which the targeted morpheme encodes both tense and agreement.

Let us assess. Of these three languages, it is clear for English that a mono-morphemic analysis is more parsimonious than a bi-morphemic one in three possible ways. Unnecessarily doubling the number of null forms is significant. For Dutch and Faroese, a bi-morphemic analysis also creates various problematic issues, though not to the extent that it does for English. It is desirable, however, to steer away from impressions that involve an overall comparison of two analyses and try to approach more precisely what the tipping point might be that triggers a switch

from a bi-morphemic to a mono-morphemic analysis. What English, Dutch and Faroese have in common is that the 3SG elsewhere form does not return in the past 3SG slot. The attempt to maintain a bi-morphemic analysis then has the consequence for English and Dutch that it requires the postulation of a contextually restricted null allomorph. What would ensure the choice for a mono-morphemic analysis is the hypothesis that an analysis postulating a contextually restricted null allomorph loses to an analysis that treats forms already postulated as competing with each other directly. Note that the contextually defined allomorph needed to maintain a bi-morphemic analysis for Faroese would be overt but this language ran into the additional problem of having to postulate homonymous forms. The hypothetical switch to a mono-morphemic system would therefore be the desire to circumvent either postulation of a contextually restricted null allomorph or a homonymous form.

Let us end this section with spelling out the consequences for pro drop. We have looked at languages with poor agreement paradigms, English, Dutch and Faroese, which, despite the fact that agreement does not express all the features that agreement expresses in Spanish and Italian, still mark some contexts with a unique form. In a contextual approach to pro drop, it is then unclear why such languages do not license null subjects in those contexts. Under the assumption that English, Dutch and Faroese are unsplit-IP languages, however, all forms that spell out agreement features also spell out tense features, as indicated in the analyses above. Hence, they run into the overspecification problem and pro drop is effectively blocked.

With this hypothesis in hand, let us now explore the morphological properties of three languages that lack pro drop despite having a rich agreement paradigm. Do they pattern with the other split-IP languages or with the other unsplit-IP languages?

3.3 Rich agreement languages without pro drop: Icelandic, Standard German and Standard French

As indicated in section 1, Icelandic is a language that shows a contrast between 1st, 2nd and 3rd person in its agreement paradigm (17). In addition, Icelandic looks like a language where tense and

agreement morphology are distinct, given that all past forms contain *-ð*, and all present tense forms do not.

(17) Icelandic present and past tense paradigms:

	present	past
1SG	-i	-ð-i
2SG	-ir	-ð-ir
3SG	-ir	-ð-i
1PL	-jum	-ð-um
2PL	-ið	-ð-uð
3PL	-a	-ð-u

Along these lines, one could postulate the following rules for tense (assuming as before that unmarked features reflect the absence of a particular feature value):

(18) *Tense*

-ð <> [T: past]

-∅ <> [T:] (present)

The singular forms could then be captured by the following rules, where the unmarked form is the elsewhere form:

(19) *-i* <> [uφ: speaker]

-ir <> [uφ:]

The problem that now arises is that the 2SG-3SG syncretism that emerges in the present tense disappears in the past tense, where it is replaced by another syncretism, namely between 1SG and 3SG. The question is how to analyze these distinct syncretic patterns in present and past tense agreement paradigms. More specifically, how can we capture the fact that *-ir*, the 3SG form in the present tense, does not reappear in the 3SG past tense slot, though it does in the 2SG slot?

The familiar options come to mind. The first one would be to induce impoverishment, a process that eliminates a feature before insertion of a morpho-phonological form. In the present tense, *-ir* appears in the 3SG context and thereby functions as the elsewhere form of the system, expressing no feature values. Hence, there is no feature value that can be impoverished such that in the past tense insertion of *-ir* is blocked in 3SG contexts. And since *-ir* is the elsewhere form in the present, *-i* is not. So even if impoverishment were possible, it would not lead to *-i* surfacing instead. The fact that in Icelandic the agreement alternation between present and past tense takes place in the 3SG again causes the problem.

But even if we were to abandon the assumption of the 3SG form being the elsewhere form and instead assume that *-i* is the elsewhere form, a bi-morphemic analysis remains problematic. In that case, the *-ir* form must be the spell out of [uφ: non-speaker]:

(20) [uφ: non-speaker] <> *-ir*

[uφ:] <> *-i*

Then, in addition, one would need to assume an impoverishment rule like (21), which refers to a second negatively marked feature ([uφ: non-addressee] or [uφ: non-participant]) in the context specification of the context, so as to avoid *-i* from appearing in 2SG past tense contexts.

(21) [uφ: non-speaker] → [uφ:] / __[uφ: non-addressee], [T: past]

Although the facts come out, the rule in (21) is highly problematic. First, there is no syntactic agreement slot where [uφ: non-addressee] could be hosted, as that slot would already be occupied by [uφ: non-speaker] itself. Moreover, there cannot be an agreement morpheme that carries the feature [uφ: non-addressee], as none of the agreement spell-out rules (i.e. the ones in (20)) refers to this feature. Hence, (21) is an impossible impoverishment rule within the Icelandic grammar. This exhausts impoverishment as a way to block *-ir* from occurring in 3SG past tense contexts.

A second way to analyze the Icelandic paradigm would be to stick with the spell-out rules in (19) and add a context-sensitive spell-out rule. One could assume that *-i* is an allomorph of *-ir*, inserted by the rule in (22).

(22) *-i* <> [uφ:] / [T: past]

Such context-sensitive rules are required for Icelandic anyway to account for the differences between the Icelandic present and past tense agreement markers in the plural (where *-jum* becomes *-um*, *-ið* becomes-*uð*, and *-a* becomes-*u*) under a bi-morphemic account. Nevertheless, (22) predicts that *-ir* is also replaced by *-i* in the 2SG past tense context, contrary to fact. A solution would be to assume that in the present tense there are two *-ir* forms, one for 2SG and one for 3SG contexts:

(23) *-ir* <> [uφ: addressee]

-ir <> [uφ:]

Although we now no longer predict *-i* to appear in the 2SG slot in the past tense, this analysis comes at a cost. First of all, the analysis now assumes two *-ir* homonyms. In addition, it also postulates two *-i* homonyms: one form that in the present tense exclusively appears in 1SG context and must be the spell out of [uφ: speaker], and a second one that appears in the 3SG slot in the past tense, inserted by (22). In sum, there are two cases of potential syncretism, one in the present and one in the past tense, and the analysis fails to capture either of them.

A third alternative to account for the agreement alternation is, again, to assume that *-ir* not only competes with *-i* in the present tense singular, but also with the past tense form *-ði*: *-ir* is not inserted in 3SG past contexts because *-ði* is inserted instead. In other words, an agreement form *-ir*, spelling out [uφ:], competes for insertion in 3SG slots with a form that also spells out a tense feature [T: past]. For competition between *-ir* and *-ði* to be possible, they must spell out a feature of the same type. The elsewhere form can therefore not be a pure agreement elsewhere, but must be a more general inflectional elsewhere spelling out both [T:] and [uφ:]. This analytical move has consequences for the rest of the paradigm, as the same analysis will carry over to other forms that

-ir competes with, such as the *-i* we see in 1SG present tense contexts. If A (*-i*) competes with B (*-ir*) and B must also compete with C (*-ði*), then by transitivity C must be able to compete with A. This way, we end up with a mono-morphemic analysis of tense and agreement, so that the spell-out rules look as in (24):

(24) *-i* <> [T:], [uφ: speaker]

-ir <> [T:], [uφ:]

-jum <> [T:], [uφ: speaker, plural]

-ið <> [T:], [uφ: addressee, plural]

-a <> [T:], [uφ: plural]

-ð*i* <> [T: past], [uφ:]

-ð*ir* <> [T: past], [uφ: addressee]

-ðum <> [T: past], [uφ: speaker, plural]

-ðuð <> [T: past], [uφ: addressee, plural]

-ðu <> [T: past], [uφ: plural]

Now, let us compare this mono-morphemic analysis with the bi-morphemic one:

(25) **-**ð <> [T: past]

-**Ø** <> [T:] (present)

-um <> [uφ: speaker, plural] / [T: past]

-jum <> [uφ: speaker, plural]

-uð <> [uφ: addressee, plural] / [T: past]

-ið <> [uφ: addressee, plural]

-u <> [uφ: plural] / [T: past]

-a <> [uφ: plural]

-ir <> [uφ: addressee] / [T: past]

-i <> [uφ: speaker]

-i <> [uφ:] / [T: past]

-ir <> [uφ:]

One may perhaps be inclined to interpret the move from a bi-morphemic to a mono-morphemic analysis as a defeat, as it gives up on a transparent analysis of agreement and tense, in which a certain number of agreement forms is listed only once and used in present and past tense contexts alike. At the same time, we again see concrete advantages of the analysis in (24). First of all, it does not postulate any homonymous forms: although the -ð consonant is part of all forms expressing past tense, the only forms that can be compared to establish the number of homonyms are the forms of which -ð is a part, and these are all distinct. Second, it induces no context-sensitive rules or impoverishment. Should one believe that rule counting matters, then the mono-morphemic analysis (10 rules) is more parsimonious than a bi-morphemic one (12 rules).

With respect to German, let's immediately zoom in on the problem that needs solving. The 3SG present tense *-t* does not return in the 3SG past tense slot.

(26) *Standard German present and past tense paradigms*

	German	
	spazieren ('to walk')	
	present	*past*
1SG	spazier-e	spazier-**te**
2SG	spazier-st	spazier-**te**-st
3SG	spazier-t	spazier-**te**
1PL	spazier-en	spazier-**te**-(e)n
2PL	spazier-t	spazier-**te**-t
3PL	spazier-en	spazier-**te**-(e)n

Since the agreement alternation happens in the 3SG contexts, impoverishment cannot be induced as a way of suppressing *-t* in the past tense, just like case of Icelandic. Under the assumption that the 3SG form is the elsewhere form, there is no feature value to delete and the required impoverishment rule cannot be formulated. This leaves inducing a context-sensitive rule as theoretical option. This rule would look as in (27):

(27) -Ø <> [uφ:] / [T: past]

This rule introduces a second elsewhere form, reserved for the past tense, and this form is phonetically zero. One might take (27) to be a necessary addition to the grammar in the absence of an alternative analysis of the data, but we have seen before that there is at least one alternative analysis, tailor-made for such data, namely a mono-morphemic one. In that vein, the absence of *-t* in the past tense can again be seen as a consequence of direct competition between the 3SG form *-t* and the form we see in 3SG past tense contexts, *-te*. For this competition to be possible, these forms must both be spell outs of the same type of morpheme, namely one that expresses both tense and agreement features. Since *-t* also competes with for instance *-e* and *-st* in the present tense singular slots, the hybrid nature of the underlying morpheme is consequently generalized across the paradigm. This gives the spell-out rules in (28):

(28) *Inflection*

-e <> [T:], [uφ: speaker]

-st <> [T:], [uφ: addressee]

-t <> [T:], [uφ:]

-en <> [T:], [uφ: plural]

-t <> [T:], [uφ: addressee, plural]

-te <> [T: past], [uφ:]

-test <> [T: past], [uφ: addressee]

-ten <> [T: past], [uφ: plural]

-tet <> [T: past], [uφ; addressee, plural]

We can now compare the two analyses and conclude that the mono-morphemic one loses to the bi-morphemic one on transparency: it is a coincidence that all forms expressing past tense have *-te*, since *-te* does not appear as a separate listing. But the mono-morphemic analysis wins in not having to postulate a null allomorph and in not having to condone an accidental syncretism. The fact that the 1SG and 3SG slots in the past look the same is immediately captured under (28) but is a coincidence under the bi-morphemic analysis, where *-te* plus *-e* are generated in the 1SG past tense slot and *-te* plus -Ø in the 3SG past tense slot. This coincidence can be circumvented by impoverishing the [speaker] feature in the past tense, so that insertion of *-e* is blocked and -Ø can be inserted in both 1SG and 3SG past tense context. But this means one would need to complicate the grammar in order to make it simpler.

Given the crux of our analysis, there is nothing inherently Germanic that would make Icelandic and Standard German mono-morphemic. This choice was based on properties of the present and past tense paradigms. Let us in this light consider the status of Standard French, a Romance language that is generally treated as a non-null subject language, as leaving out the argumental subject leads to ungrammaticality (example from Sheehan 2016):[7]

(29) *(Tu) chantes bien *Standard French*

You sing.2SG well

"You sing well"

Let us therefore see if we can relate the lack of pro drop to morphological properties. Consider the following paradigms:[8]

(30) The present and imperfect tense agreement paradigms in Standard French

[7] There is strong evidence supporting the view that in Colloquial French subject clitics act as agreement markers (cf. Koeneman & Zeijlstra 2014 for an overview of the literature). Under such an analysis, Colloquial French is a full-fledged pro drop language. This is expected under our analysis, since the preverbal clitics justify the postulation of an agreement morpheme separate from the morpheme expressing tense.

[8] We can leave out discussion of the third conjugation as it does not form a natural conjugation class but rather a set of irregular verbs (including the so-called *-re* verbs).

	First conjugation (*-er* verbs)		Second conjugation (*-ir* verbs)	
	present	imparfait	present	imparfait
1SG	demande /dəmɑ̃d(ə)/	demandais /dəmɑ̃dɛ/	choisis /ʃwazi/	choisissais /ʃwazisɛ/
2SG	demandes /dəmɑ̃d(ə)/	demandais /dəmɑ̃dɛ/	choisis /ʃwazi/	choisissais /ʃwazisɛ/
3SG	demande /dəmɑ̃d(ə)/	demandait /dəmɑ̃dɛ/	choisit /ʃwazi/	choisissait /ʃwazisɛ/
1PL	demandons /dəmɑ̃dɔ̃/	demandions /dəmɑ̃djɔ̃/	choisissons /ʃwazisɔ̃/	choisissions /ʃwazisjɔ̃/
2PL	demandez /dəmɑ̃de/	demandiez /dəmɑ̃dje/	choisissez /ʃwazise/	choisissiez /ʃwazisje/
3PL	demandent /dəmɑ̃d(ə)/	demandent /dəmɑ̃dɛ/	choisissent /ʃwazis/	choisissaient /ʃwazisɛ/

The lack of pro drop can be related to a property it shares with Icelandic and Standard German. The singular form in the first conjugation is not null but /-ə/. If it were null, we would expect the stem-final *-d* of *demand* to be dropped in pronunciation (like in *quand* /kɑ̃/ 'when'), contrary to fact. There is no evidence, however, that this /-ə/returns in the singular part of the imparfait. In order to account for its absence there, we cannot induce an impoverishment rule (as /-ə/ is the elsewhere form) but we must postulate a null allomorph. We therefore run into the exact same issue as Standard German. To circumvent postulation of a null allomorph, a mono-morphemic analysis of tense and agreement can be assumed instead.

There is in principle a way to maintain a bi-morphemic analysis, though, and not having to postulate a null allomorph at the same time. One could assume that /-ə/in the present tense is not an agreement form but a (present) tense form. In that event, one could postulate a null elsewhere

agreement form that appears in the singular, contrasting with 1PL -/õ/ and 2PL -/e/. Under such an analysis, -ɛ in the past tense can be analyzed as the past tense marker followed by the same null agreement form that appears in the present tense. In this way, the spell-out forms can be taken as evidence for the existence of separate tense and agreement morphemes. The problem with this analysis, however, is that it only looks at the singular part of the paradigm. The main reason for assuming that a particular form expresses (present) tense (here, /-ə/) is that the form is maintained when you change the agreement form. But that never happens in Standard French, as the sjwa does not co-occur with -/õ/ and -/e/. Likewise, the main reason for assuming that a form expresses past tense (here, -/ɛ/) is that this form is retained when you change the agreement form, but again this never happens in French, as -/ɛ/ does not co-occur with -/õ/ and -/e/. Put differently, it is a coincidence under this analysis that every time the agreement form becomes overt (as in 1PL and 2PL contexts) the present tense form becomes covert. And it is a coincidence that every time the agreement marker becomes overt in the past tense (as in 1PL and 2PL contexts) the elsewhere past tense form is suppressed in favor of *-i-*.

The bi-morphemic analysis is suspicious and must be maintained at the expense of ignoring what look like straightforward complementary distribution effects between overt forms. It is not the case that /-ə/ remains absent in 1PL and 2PL present tense contexts because it competes with a null tense marker that is realized in these contexts but because it competes with -/õ/ and -/e/ directly. Instead of taking -/ɛ/ in the past tense to compete with *-i-*, which in turn requires that a null agreement form competes with -/õ/ and -/e/, the form -/ɛ/ can be taken to directly compete with -/jõ/ and -/ie/, explaining its absence in 1Pl and 2PL contexts.

The second conjugation does not provide phonological evidence for an underlying sjwa in the singular. The 1SG form of *choisir* is /ʃwazi/, where the final /s/ is dropped in pronunciation but would have been retained under the presence of a sjwa (like the /d/ in /dəmɑ̃d/. This suggests

that the elsewhere tense/agreement form is a contextual null allomorph, so that we need the spell-out rules in (31):

(31) -/ə/ <> [uφ:] / [T: past] / Conjugation 1

-Ø <> [uφ:] / [T: past] / Conjugation 2

One may wonder if allowing this null allomorph would not mean that we have to allow it in Standard German as well. Note, however, that there is still a fundamental difference between postulation of the contextual null form in Standard French and the one in Standard German. In the latter case, an alternative analysis was available based on direct competition of overt forms: by letting the elsewhere form that is not context-sensitive (the 3SG-*t* we see in the present tense) compete directly with the form we find in the past (-*te*), the null allomorph is avoided (see (28)). The strategy to let other forms compete directly so as to avoid a null allomorph does not work for Standard French for the simple reason that the elsewhere tense/agreement form used in the first conjugation (the - /ə/) has nothing to compete with in the second conjugation paradigm if it wasn't for the hypothesized null allomorph. In the absence of an alternative analysis that avoids a null allomorph, adopting one is unproblematic.[9]

To summarize the discussion about Icelandic, Standard German and Standard French, we conclude that neither a bi-morphemic nor a mono-morphemic analysis can be discarded off the bat, but that a mono-morphemic analyses appears to have a number of advantages, at least *prima facie*. In the next section, we will look in more detail at what needs to be theoretically assumed exactly in order to maintain a connection between bi-morphemicity/mono-morphemicity and pro drop/no pro drop that is empirically sound, i.e. correctly captures the languages under discussion.

[9] Strictly speaking, it would be possible to prevent postulation of the null allomorph in (31) by letting -/ə/ compete directly with the stem form of the second conjugation verb, i.e. /ʃwazi/. This would require conflating the Agr, T and V positions in the second conjugation, but conflating only Agr and T in the first conjugation. However, conflation of Agr, T and V is out for semantic reasons: tense must have scope over the predicate denoted by the VP (cf. von Stechow 2003).

4 The role of acquisition in explaining the distribution of pro drop

We have hypothesized that a bi-morphemic analysis of tense and agreement is a prerequisite for the licensing of null subjects. In order to test this hypothesis, one should look at the tense and agreement paradigms of individual languages and determine for each of them what the most adequate analysis is. This was the topic of the previous section. The inherent problem of this exercise, however, is that there is not one way in which an individual language can be analyzed. In fact, every language we have looked at can in principle be analyzed in a bi- or mono-morphemic way. Our theoretical toolkit is rich enough for that. What we should do, therefore, is compare a bi- and mono-morphemic analysis for each language and establish which of the two is the most parsimonious one. Although this sounds straightforward in theory, in practice it is not.

Take for instance our discussion of Icelandic. One might conclude from it that there are at least three reasons for taking a mono-morphemic analysis to be the most parsimonious one: (i) it postulates fewer rules, (ii) it postulates fewer null forms, and (iii) it postulates fewer homonyms. If this was all there was to it, the conclusion would be simple. However, there is also an obvious argument for analyzing Icelandic bi-morphemically. Under a mono-morphemic analysis, the fact that all forms spelling out past tense start with a *-ð* must be understood as a diachronic left-over of times in which *-ð* was the spell out of a separate tense morpheme. Synchronically, however, this fact is an accident not captured by the analysis, as it does not postulate a vocabulary entry *-ð*. One could therefore also argue that a bi-morphemic analysis is more parsimonious than a mono-morphemic one in that it captures a particular morphological regularity in a straightforward way. One could of course argue that restricting rules, null forms and homonyms is more important than capturing this regularity, but we see no intrinsic reason why this would be so. The attempt of relating pro drop to bi-morphemicity and lack of pro drop to mono-morphemicity would therefore rest on a stipulation.

As a way out of this conundrum, we take it to be helpful to approach this issue from the perspective of language acquisition, as this shifts the basis on which the choice between two analyses for a particular language is plausibly made to the language learner, who ultimately has to carry out this task. To see this, consider again the two analyses for Icelandic. We have seen that a bi-morphemic analysis needs to postulate twelve rules whereas a mono-morphemic analysis can make do with just ten. One could therefore conclude that the mono-morphemic analysis is more parsimonious from a theoretical point of view. But if the choice between these two analyses is made by the language-learning child, it becomes extremely unlikely that rule counting would make the difference here. From an acquisition perspective, we can fairly safely conclude that rule counting is irrelevant.

Likewise, one may wonder how plausible it is to assume that the child's job is to maintain morpho-syntactic transparency at all costs, i.e. to assume split projections for tense and agreement even if the input data were to show a 100% fusional morphology for these properties in all paradigms. That would mean that morphological data should have no effect on the morpho-syntactic representation and that the assumed syntactic tree must therefore be postulated as innate. It is clearly beyond the purpose of this paper to have a principled discussion about the pros and cons of such a "universal base" approach to syntax. Note, however, that such an approach makes it harder in general to account for correlations between syntax and morphology which previous studies have shown to exist in their own right, and of which Bobaljik & Thráinsson (1998) is a prime example (see also Holmberg & Roberts 2011, pace Newmeyer 2004). It is explicitly the purpose of this paper to try and get significant mileage out of relating morphology to the underlying representation in a more direct way, namely a better understanding of the pro drop phenomenon.

This leaves null forms and homonyms as the two left-over properties that play a role in the analyses of the last section, where they served as potential last resort mechanism to keep an analysis of tense and agreement bi-morphemic. Let us therefore determine which role they play

exactly in the acquisition of morphological systems. Now, it is generally agreed upon that direct competition guides the native language learner in distinguishing different forms that belong to the same class in the input (Pinker 1984 and subsequent work). This entails that direct competition is the driving force for the acquisition of the relevant spell-out forms and the morpho-syntactic features that they externalize. The importance of this is that it puts the role of null forms and homonyms in an acquisitional perspective. When can a child straightforwardly postulate a null form or homonym? Let us entertain the following hypotheses:

(32) Postulation of a null form happens if there is no other form in the grammar that spells out the same feature set.

(33) Postulation of two homonymous forms happens if there is no feature set that can be spelled out by just one form.

Let us discuss each in turn.

(32) entails that a straightforward null form fills a gap in the paradigm by contrasting with an overt form. In that case, the null form can be said to be paradigmatically licensed. An example of this would be the null form that spells out present tense, [T:], in all the languages under discussion. In Italian such a null form contrasts with the imperfect *-v* marker, in English with the past tense marker *-ed*, etc. In either case, there is no other form already spelling out [T:]. Another example would be the postulation of a null form in English for the 1SG and 2SG present tense contexts. This form contrasts with *-s* and spells out a feature, [uφ: participant], that no other form spells out already (*-s* spells out [uφ:]).

(33) entails that postulation of a homonym is what a child may have to do if two identical forms share no plausible feature set. If they did, one form spelling out the shared features would suffice. In other words, the child will try to maximize syncretisms but will not always succeed. An example of a straightforward homonymous pair is the *-t* form that shows up in 3SG and 2PL present tense contexts in German (cf. (26)). Since these contexts involve different person and number features, the two *-t* forms are only accidentally alike, and analyzing them as homonyms is

straightforward. Note also that an analysis that takes both *-t* forms to be the same (Frampton 2002; Müller 2006) has to still account for the fact that the 3SG *-t* does not return in the past tense whereas the 2PL *-t* does. Since we have seen how problematic it is to restrict impoverishment to 3SG contexts, this contrast provides confirmation that the two *-t* forms are accidentally homonymous.

By contrast, the 1st and 3rd person plural forms *-en* in German, cannot be taken to be homonymous as they spell out the same feature set [uφ: plural] (with plural *-t* then spelling out [uφ: plural; addressee]).

The existence of straightforward null forms and homonyms presumes the existence of non-straightforward ones, to which we will now turn. Recall that the analysis for English has to capture the fact that the present tense *-s* does not return in the past tense. One way to account for this is to assume the existence of a null allomorph in the past tense that blocks the *-s* from being inserted. Whereas *-s* is the elsewhere agreement form in the present tense, a null form does that in the past:

(34) *-s* <> [uφ:]

(35) -Ø <> [uφ:] / [T: past]

This null form spells out the same feature set as *-s* does (note that [T:] in (35) merely describes the context in which the form does that), so this constitutes a non-straightforward null form. Now, rather than putting a ban on null allomorphs altogether, we propose that at this acquisition step the child first assumes that *-s* competes with the past tense form *-ed* directly and that these forms must therefore target the same morpheme. In other words, the learner sees an analytical alternative, and one that makes sense from an acquisitional point of view: The child will rather account for the complementary distribution of *-s* and *-ed* by assuming direct competition between forms that it has direct evidence for (because they are overt) than create a competition by postulating a form that it has no direct evidence for because it is phonologically null. The consequence of this

analytical choice is of course that the English agreement paradigm must be mono-morphemic and that consequently English will lack pro drop. This analysis extends to Standard German, Dutch and Faroese.

The same point can be made about non-straightforward homonyms. Take the case of Icelandic. The fact that the 3SG present tense form *-ir* does not return in the 3SG past tense can be captured by assuming that the *-i* form appearing there is an allomorph of *-ir*. So again we have two elsewhere forms, one context-specific.

(36) *-ir* <> [uφ:]

(37) *-i* <> [uφ:] / [T: past]

Since the *-i* also occurs in the 1SG of the past (cf. (17)) and both *-i* forms sit in the singular part of the paradigm, they share a feature. Therefore, this analysis creates a homonym that is not straightforward. One could of course assume that the same *-i* is inserted in 1SG and 3SG past contexts (i.e. treat these forms as syncretic). However, not only would this require an impoverishment rule that deletes the [speaker] value of the past singular morpheme, it also creates a non-straightforward homonym with the 1SG form that occurs in the present tense, which is also *-i*. Since the 1SG singular and 1SG past form obviously share number and person features, this homonym would also be non-straightforward. By hypothesis, then, the learner assumes that there is direct competition between 3SG present *-ir* and 3SG past *-ði* because alternatively it would have to postulate a non-straightforward homonym. The consequence of that acquisition step is that the Icelandic agreement paradigm becomes mono-morphemic and that Icelandic, like Standard English, must lack pro drop, which is indeed the right result.

It is important to note that the proposed restrictions are not hard ones. We do not put a ban on null forms and homonyms and even condone non-straightforward ones: (32) and (33) are meant to reflect learner strategies, they are not intended as inviolable constraints of the grammar. To see this, take a look again at the elsewhere forms we postulated for Standard French in (31), repeated below as (38):

(38) -/ə/ <> [uφ:] / [T: past]

-Ø <> [uφ:] / [T: past]/ Conjugation 2

Given (32), this null allomorph constitutes a non-straightforward one. The fact that it still ends up as part of the grammar, in contrast to the null allomorph in English, Dutch and Standard German, is the lack of an alternative analysis that circumvents its postulation. Whereas the 3SG elsewhere forms in the Germanic languages can be taken to directly compete with the past tense forms, there is nothing inflectional in the paradigm of the second conjugation that -/ə/ can compete with. This opens the door for the null allomorph.

5. Discussion

In the previous section, we have made explicit what needs to be assumed in order to maintain a correlation between the bi-morphemic realization of tense and agreement and pro drop. We can now summarize our proposal on the distribution of argumental pro drop as the combination of a grammatical claim and an acquisitional claim, formulated in (39) and (40), respectively:

(39) In order to license an empty argumental subject, the licenser cannot be underspecified, nor overspecified by features that are semantically incompatible with those of an argumental subject.

(40) In order to account for a complementary distribution effect between A and B in a paradigm, direct competition of A and B takes priority over an analysis that explains the effect by postulating a (non-straightforward) null form or homonym.

Put differently, postulation of non-straightforward null forms and homonyms is a last resort strategy only employed if there is no viable alternative, and we propose that a mono-morphemic analysis counts as a viable alternative to a bi-morphemic one. A consequence of this is that the assumption of a transparent relation between morpho-phonological forms and the morphemes that they spell out may be a default assumption held by the learner but is not an acquisitional goal in itself, as it may clash with other acquisitional strategies.

Another consequence is that paradigmatic effects are achieved without reference to the paradigm itself. If in English *-s* and *-ed* compete directly, there must be an underlying morpheme that they both target, expressing tense and agreement information. But if *-s* is also to compete with for instance the null spell out in 1Sg and 2SG contexts, the form -Ø must be the spell out of the same morpheme, etcetera. The consequence is that all contexts have a syntactic representation with an unsplit IP in the terminology of Bobaljik & Thráinsson (1998). Hence, English lacks pro drop across the board. The same is true in Icelandic and Standard German, languages that in principle have forms with enough agreement features to license a null subject if it wasn't for the overspecification problem.

With this analysis in place, we would like to discuss two prominent issues that now ensue:

1. What causes the overspecification problem? Why would the presence of a tense feature render pro drop impossible?
2. If a consequence of being mono-morphemic is that a language lacks pro drop across the board, how do we account for the existence of partial pro drop varieties, where pro drop is only possible in certain contexts?

In the next sections, we will turn to these questions.

5.1 The overspecification problem

Under our analysis of pro drop, the licensing morpheme cannot be underspecified or overspecified. The underspecification part is a very prominent feature in the pro drop literature, especially since Rizzi (1982), and an explicit or implicit assumption in many accounts: an empty subject must be reconstructed and therefore the agreement system of a language must be rich. We agree with this general intuition and only take issue with the paradigmatic formulation of it, for reasons explained in the introduction.

The overspecification problem is new, however, and therefore requires some more discussion. We argued that a morpheme that expresses both tense and agreement features loses

the option of licensing a null subject because the tense features it carries are incompatible with the interpretation of the null subject that it is supposed to license. There is an additional question, however: why can the tense features on the agreement marker not simply be ignored by the null subject?

As a first step towards an answer, let us assume that an empty subject is a 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*. This means that maximal syntactic uniformity between pro-drop and non-pro-drop languages can be upheld.

Under the view that pro drop involves the presence of a covert subject *pro*, a correlation between the richness of agreement and pro drop also naturally follows. As there is only one *pro*, it must be featurally underspecified in the lexicon. After all, enriching the number of *pro*s in the lexicon to at least six (one for every person/number context) would be conceptually very awkward (cf. Holmberg 2005). 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.

To give an example, the uninterpretable 1SG φ-feature values on the agreement marker in the Italian example in (41) value *pro*'s unvalued φ-feature, as shown in (42). At LF, *pro* is consequently interpreted as the 1st person singular pronominal subject.

(41) Canto Italian

Sing.1SG

'I sing'

(42) Before valuation: $[_{AgrP}\ [_{DP}\ [\varphi:\ \underline{\quad\quad}]$ ... Agr [uφ: speaker]]

After valuation: $[_{AgrP}\ [_{DP}\ [\varphi:$ speaker] ... Agr [uφ: speaker]]

Such an analysis straightforwardly predicts that the agreement paradigm must be rich, as otherwise *pro* can never receive the same interpretation as the corresponding overt pronouns (see also Saab 2008, forthcoming, who argues that for the subject to be left unpronounced, featural identity is required). If pronominal paradigms are rich, the agreement paradigm should be equally rich for *pro* to be interpreted as a subject.[10]

It now straightforwardly follows that pro drop is licensed in languages in which tense and rich agreement morphology is bi-morphemic and the relevant features are hosted on two different functional heads, like in the Romance languages. We have yet to explain, though, why in languages like Icelandic and Standard German, in which tense and agreement are expressed mono-morphemically and the relevant features must be 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, as in (42) (again see also Saab 2008)?

The answer, we argue, has to do with the fact that non-pro drop languages have an unsplit-IP. This means that the relevant morpheme is I instead of T or Agr. The essence of a category that conflates T and Agr is that it reduces the status of T and Agr to that of values of the categorial feature I. In other words, if T and Agr were additional features on I, we would not have a conflated

[10] The idea that the agreement marker values the features of the subject has been criticized by Holmberg (2005, 2010) and Roberts (2010), who argue that agreement features are not only uninterpretable on F, but also unvalued. In this, they follow Chomsky (2000, 2001), who argues that only lexically valued features are interpretable features. If *pro* is an empty subject, however, pro drop cannot be treated along the same lines. In its traditional conception, *pro* needs to be valued by rich agreement, and not the other way around. If only interpretable elements can value unvalued features, such as the ones underlying agreement morphology, and *pro* is by definition unvalued, there can be no inherently valued φ-features present in the tree that can value *pro*.

However, this criticism against uninterpretable features valuing interpretable ones is currently not grounded in existing theories of agreement anymore. Following Pesetsky & Torrego (2007) checking and valuation are nowadays taken to be operationally distinct (see also Arregi & Nevins 2012 and Bjorkman & Zeijlstra 2019). Uninterpretable features must be checked; unvalued features must be valued, and nothing forbids interpretable, unvalued features or valued, uninterpretable features (cf. Pesetsky & Torrego 2007). Hence, there is no theoretical objection against uninterpretable features valuing *pro*.

category to begin with. The fact that T and Agr are values of I has non-trivial consequences for the content of the agreement subfeature. Since tense is an interpretable value, I itself must be interpretable. But this means that the agreement value 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 values would feed interpretation at LF, which would crash the sentence: tense and φ-values cannot be interpreted in the same semantic position, as part of the same feature I, because they are of different semantic types. In other words, no coherent interpretation for the morpheme I is possible.

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 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 in the first place. But if tense on I must be valued, φ cannot be. Hence, I-features that comprise both tense and φ-values must be valued in the course of derivation. To come back to the Icelandic example, [$_I$ T: past; φ: speaker, plural] cannot be a lexical item for reasons we have just seen, but [$_I$ T: past; φ: _,_] can be. Even though the φ-values on I are strictly speaking part of an interpretable feature, they never feeds LF because they cannot be valued next to tense and must therefore stay unvalued.

Now, the only way left to value [φ: ,] on the I-head is through agreement with the subject. This consequently requires of the DP subject that its φ-features are valued. Since *pro* is inherently unvalued, it is an impossible subject to have. It would leave agreement unvalued and leave *pro* unvalued at the same time. This, then, rules out pro drop in a language where agreement is a value of I.

The difference between pro-drop languages and non-pro-drop languages, then, is that in the former, agreement markers are inherently φ-valued members coming from the lexicon, whereas in the latter they can only be φ-valued in the course of the derivation. The distinction

between pro-drop and non-pro-drop languages thus reduces to the traditional distinction between inherent vs. contextual agreement morphology.

5.2 Partial pro drop

Above, we have shown how Germanic languages can lack pro drop across the board without making reference to the paradigm as a whole. It is not because the paradigm of for instance English is poor overall that it lacks pro drop. It is because the spell-out forms postulated require the presence of an underlying morpheme that is overspecified for the purposes of licensing a null subject. Our approach is therefore compatible with a contextual approach to pro drop in which the possibility of licensing a null subject is determined per context. The outcome of this will be the same for every context in English, however, due to the overspecification property. If the morphological input either leads to a syntax with a split- or an unsplit-IP and languages are predicted to either have or lack pr drop because of this choice, the existence of partial pro drop varieties comes as a surprise. Let us look at some partial pro drop that are typologically related to the Germanic languages that we have discussed.[11]

Pro drop in Germanic varieties mostly occurs in 2nd person singular contexts, as examples in (43) from Bavarian (cf. Bayer 1984) and (44) from Frisian (cf. De Haan 1984) illustrate, but pro drop in 1st and 2nd person plural are also attested in some Bavarian varieties as shown in (45).

(43) a. Kumm-st (du) noch Minga, dann muas-st *pro* me bsuacha
Come.2SG (you) to Munich, then must.2SG me visit
'If you come to Munich, you must visit me'.

b. Ob-st (du) noch Minga kumm-st, …
if.2SG (you) to Munich come.2SG
'If you come to Munich, ...'

[11] This means that we currently have nothing to say about Standard Finnish and Hebrew.

(44) a. Miskien moat-st (do) my helpe
Perhaps must.2SG (you) me help
'Perhaps you should help me.'

b. Ik denk dat-st (do) my helpe moatst
I think that.2SG me help must
'I think that you should help me.'

(45) a. Fahr-ma (mir) noch Minga?
drive.1PL (we) to Munich
'Do we drive to Munich?'

b. Ob-ts (es/ihr) noch Minga kumm-ts,…
whether.2PL (you.PL) to Munich come.2PL
'Whether you come to Munich, …'

The question is how partial pro drop is licensed in these varieties but not in Standard German or Dutch.

Bavarian German varieties have an agreement paradigm with a similar structure as that of Standard German but they famously lack a past tense (cf. (47), constructed on the basis of the information in Fuß 2005). To express a past tense, a compound tense is used instead. One might be inclined to take this as relevant: if there is no past tense, there is no evidence that the 3SG *-t* disappears in the past tense, as there is in Standard German, and there is therefore no trigger to switch to a mono-morphemic system. Taking the lack of a past tense to be the relevant factor is problematic for two reasons, though. First of all, Frisian does have a past tense, one in which the 3SG *-t* remains absent in the 3SG context (cf. (47)). This means that such an account would only work for Bavarian dialects. Second, the absence of a past tense does not entail that the language will have a separate agreement morpheme. For semantic reasons, Bavarian dialects must have a representation for present tense. This tense is never overtly expressed, however. Assuming that [T:] is spelled out by a null form would be unproblematic if there existed a form spelling out [T:

past] next to a form spelling out [T:], as is the case in Standard German. Since such a form is absent in Bavarian, this would require postulation of what we called a non-straightforward null form. This assumption can be avoided by assuming that the overt agreement forms also spell out present tense. This then gives rise to spell-out rules such as (46), where *-e* is the vocabulary item inserted in a morpheme that expresses both tense and agreement.

(46) *-e* <> [T:], [uφ: speaker]

In sum, Bavarian dialects are mono-morphemic in the same way as Standard German, despite (and in fact because of) the lack of a past tense.

(47) Bavarian and Frisian (present and past) agreement paradigms

	Bavarian		Frisian	
	present	past	present	past
1SG	-Ø	-	-Ø	-te
2SG	-st	-	-st	-test
3SG	-t	-	-t	-te/*-tet
1PL	-an	-	-e	-ten
2PL	-ts	-	-e	-ten
3PL	-an	-	-e	-ten

This means that, in principle, the Bavarian and Frisian agreement paradigms should be mono-morphemic. But as explained before, this does not necessitate that all contexts, i.e. all cells in the paradigm must be so. If there is on the one hand evidence for mono-morphemicity but on the other hand there are phenomena in the language that call for particular agreement forms to be analyzed differently, the language learning child could converge at a mixed paradigm, one that is bi-morphemic in some specific contexts and mono-morphemic in all others. However, for such a mixed paradigm, the language learning child must indeed be confronted with evidence that either a fully monomorphemic or a fully bimorphemic analysis of the paradigm cannot be upheld.

We argue that this indeed is what underlies partial pro drop in both Bavarian dialects and Frisian. Concretely, we relate this to the fact that all these varieties have complementizer agreement in the relevant, pro-drop licensing contexts (Fuß 2005). Rosenkvist (2009:163) states that it is exceptionlessly the case that pro drop in Bavarian is only allowed in person/number contexts that allow complementizer agreement. Since the 2SG *-st* marker in Bavarian and Frisian also shows up on the verb in main clause straight orders, in which the finite verb follows the subject ('du kumm-st' (you come.2SG), 'do moat-st' (you must.2SG), etc.), it is naturally analyzed as an agreement ending when it appears on the complementizer, as illustrated in the b-examples in (43)-(45). The fact that the subject (*du* or *do*) does not show up obligatorily after the finite verb or complementizer can then be construed as an argument in favor of these specific contexts being bi-morphemic and, therefore, pro drop may occur in these specific contexts. The lack of partial pro drop in Standard German can then be related to the lack of complementizer agreement.[12]

Identifying the relevant factor is not the same as explaining why that factor is relevant. Two questions arise:

(i) Why would complementizer agreement, which basically introduces the same affix in the clause a second time, be related to partial pro drop?

(ii) Why is partial pro drop possible in Frisian and Bavarian main clauses, where no complementizer occurs?

The two questions yield a paradox: (i) suggests that complementizer agreement is essential in the licensing of partial pro drop, whereas (ii) suggests that it cannot be.

From the perspective of our analysis, this paradox can be resolved as follows. Complementizer agreement creates morphemic transparency and shows the learner an environment in which an agreement form clearly occurs in the absence of other inflectional

[12] Van Alem (2020; forthcoming) observes for Frisian that the inflected complementizer must be adjacent to the subject, which is not a requirement for agreement in general. She therefore concludes that the complementizer does not carry an agreement marker but a subject clitic. If so, Frisian does not have partial pro drop and there is nothing to explain. In Bavarian, however, there is no such adjacency requirement and therefore this variety is more readily analyzed as a partial pro drop language.

features. After all, (Germanic) inflected complementizers do not show morphological alternations caused by tense properties expressed elsewhere in the clause, revealing that complementizer agreement markers cannot be exponents of tense/aspect features. Complementizer agreement therefore provides evidence for the existence of a morpheme expressing only agreement. This enables the learner of such varieties to generalize over e.g. the *-st* form occurring on V and the complementizer by adopting the spell-out rule in (48):

(48) *-st* <> [uφ: addressee, singular]

Complementizer agreement in a particular person/number context then triggers a syntactic analysis for that context with separate tense and agreement heads. Since 2SG *-st*, as well as 1PL *-ma* and 2PL *-ts*, are featurally neither under- nor overspecified, they will consequently license pro drop straightforwardly. The language learner adopts a mono-morphemic analysis for the variety it is acquiring and only overrides this analysis by a bi-morphemic one in contexts that provide positive evidence for a split-IP system.

In sum, the complementizer agreement input leads to a specific bi-morphemic analysis of the context in which it appears, while the mono-morphemic analyses are retained for the other person/number contexts.[13] Consequently, pro drop is licensed in all person/number contexts in which complementizer agreement appears (e.g. in all 2SG contexts), including main clauses, where no complementizer is present. It is not the presence of the complementizer as such that matters but rather the fact that the syntactic representation in that relevant context only has separate tense and agreement morphemes.[14]

---

[13] An alternative possibility is to assume that these varieties have a complete, uniform mono-morphemic paradigm but the learner adds a separate agreement (and tense) morpheme to the lexicon. This gives two ways of realizing a clause for the 2nd person singular, in e.g. the 2SG context, bi-morphemically (which allows pro drop) and mono-morphemically (which does not allow pro drop).

[14] It should be noted that complementizer agreement is not always obligatory. In Zurich German (Cooper & Engdahl 1989; Cooper 1994, see also Weiß 2005 for discussion), complementizer agreement is optional, as shown in (1):

(1) a. Ich ha nöd gwüsst dass in Züri wohnsch — Zurich German
I have not known that __ in Zurich lives

b. Ich ha nöd gwüsst dass-t in Züri wohnsch
I have not known that.2SG in Zurich lives
'I did not know that you live in Zurich

It is unfortunately not the case that partial pro drop is only claimed to exist in varieties with obligatory or optional complementizer agreement, and counter examples exist. Here, we refer to Swabian, an Alemannic variety where no overt 2SG subject has to occur in the absence of an overt marker on the complementizer (Haag-Merz 2006; Bohnacker 2013):

(49) weil wenn Ø daa runder kucksch, des isch [total!] [tief!] Swabian
because if __there down look.2SG it is totally deep
'…because if you look down there, it's totally deep.'

Bohnacker notes that there is an interesting asymmetry here, though: 2SG pronouns are omitted significantly more often after an inverted verb than after a complementizer (71% versus 14% of 2SG contexts).[15] This may suggest that what looks like pro drop here is actually a different, more local phenomenon. The reason is that 'real' pro drop involves licensing of a null subject as a syntactic category, whereas the data in Swabian are more compatible with the fact that an adjacency relation between an agreement affix and a subject pronoun immediately following is particularly amenable to processes of phonological reduction in spontaneous speech (which in fact is where Bohnacker draws her data from). Such phonological reduction processes can be observed in Standard German as well, particularly in fast speech (cf. Trutkowski 2016; thanks to Eric Fuß for pointing this out to us). We leave it to future research to establish whether the rate of pro drop in main and embedded clauses can be used to make an analytical choice between pro drop or phonological reduction more generally.

6. Conclusion

Despite the fact that pro drop is one of the better studied phenomena in linguistics, the question why pro drop is so pervasive in Romance and so unpervasive in Germanic has resisted a principled

.

[15] For straight orders, with the subject in first and finite verb in second position, the percentage of pro drop cannot be established as it interferes with topic drop, an independent phenomenon.

answer. It is generally agreed upon that agreement somehow plays a role but, as we have shown, both a paradigmatic and a contextual approach towards capturing this correlation run into problems. This must mean that at least one of the assumptions generally employed, either explicitly or implicitly, could be wrong. The job is then to identify it.

In this paper, we have proposed that part of the explanatory burden should be placed outside of the grammar, namely in the acquisition process. Concretely, we argued that bi-morphemic expression of tense and agreement is not an analytical goal of the learner but at most a default assumption in the learning process. Second, we argued that the postulation of a null morpheme or homonym is constrained and will only be entertained as an option by the child if a viable alternative is not available. Together, these assumptions create the grammars we see in Romance and Germanic, with the fundamental pro drop difference captured: Only bi-morphemic expression of tense and agreement can license pro drop, provided that the agreement is rich enough. Although these assumptions obviously have repercussions, we believe that they are nevertheless reasonable to make. The fact that morphological intransparency exists means that there must generally be an end to a fully transparent syntax-morphology mapping. Adopting null forms to uphold morphological transparency causes friction with the fact that from a learnability perspective they require clear evidence.

Given the vast literature on pro drop, there are many issues that we have skimmed over or not even addressed. In the remained of the paper, we bring up a few of these.

For one, we have deliberately restricted ourselves to discussing a small number of Germanic and Romance varieties. The reason for this is that only by comparing small differences between relatively closely related languages can one come up with the fine-grained analyses pursued above that subsequently lead to a new hypothesis. Needless to say, such a hypothesis should be subject to typological evaluation, a project that will be subject of further research.

Apart from this, much more can be said about diachronic developments. Both Old Norse and Old High German were pro-drop languages, but Icelandic and Standard German lost this

property. In principle, the loss of pro drop could be the result of a small change in the agreement patterns of these languages in our analysis, causing a switch from a bi- to mono-morphemic analysis, which leads to a catastrophic change. If Standard German were to adopt an overt allomorph for the 3SG past context, we predict it would immediately become a pro-drop language. A contextual approach to pro drop also allows a more fine-grained development, in contrast to a paradigmatic approach. We have seen that there can be circumstances under which pro drop is partially possible in a non-pro drop language. This means that pro drop can be partially retained in specific contexts if a language switches from a bi-morphemic to a mono-morphemic analysis of tense and agreement. A proper understanding of these diachronic changes in the realm of pro drop, and one that gets the timing right, requires a careful look at the data, which exceeds the purposes of the present paper (cf. Wratil 2011 for some further discussion).

Another area where future research is called for is the acquisition of pro drop. We argue that pro drop can only be acquired when children start to entertain the hypothesis that the agreement pattern of their mother tongue is bi-morphemic. As we take featural transparency as a starting point (children only converge to featurally intransparant mono-morphemic agreement paradigms if bi-morphemicity turns out to be only learnable by postulating particular instances if null forms or homonyms), we predict that children acquiring Icelandic, Standard German or French temporarily go through a pro drop stage before acquiring that the language lacks it. This should of course be subject of further study. We also make very precise predictions for the acquisition of partial pro drop. For instance, as pointed out by Theresa Biberauer (p.c.), children can only understand partial pro drop in Bavarian in our analysis after having acquired complementizer agreement. Again, this should be subject of further study.

We should point out, as we already did in section 2, that what we have discussed in this paper is (partial) pro drop that is always available for specific person and number contexts. There is nothing in our analysis that makes this the only type of partial pro drop possible. For one, there

are partial pro-drop languages like Finnish, Brazilian Portuguese and Hebrew, where pro drop seems always possible for 1st and 2nd person and not for 3rd person, as shown for Finnish below.

(50) a. (Minä) puhun englantia
I speak.1SG English

b. (Sinä) puhut englantia
You=SG speak.2SG English

c. *(Hän) puhuu englantia.
He speak.3SG English

d. (Me) puhumme englantia
We speak.1PL English

e. (Te) puhutte englantia
You=PL speak.2PL English

f. *(Te) puhutte englantia
They speak.3PL English

However, a 3rd person subject pronoun can be null when it is bound by a higher argument, as the following examples by Holmberg (2005) show:

(51) a. Se oli Tarjallei pettymys [ettei häni,j/ $\varnothing_{i/*j}$ saanut lukea latinaa koulussa].
it was Tarja-ALL disappointment that-not she could study Latin school-INE
'It was a disappointment for Tarja that she couldn't study Latin at school.'

b. Poikieni mielestä oli noloa kun hei,j/ $\varnothing_{i/*j}$ jäivät kilpailussa viimeiseksi.
boys-GEN opinion-ABL was embarrassing when they came race-INE last
'The boys found it embarrassing when they came last in the race.'

This shows that what underlies Finnish pro drop is a different phenomenon from the one we described and, following Holmberg & Sheehan (2010), arguably involves Control instead of *pro*. The fact that in these languages the distribution of PRO is substantially richer than in the languages we discuss may naturally have consequences for the existence and/or distribution of *pro*. If so, these languages are outside the domain of application of our proposal.

Finally, we note that the phenomenon known as the radical pro-drop languages, where every argumental pronoun can be dropped (cf. Huang 1984, Jaeggli & Safir 1989, Neeleman & Szendrői 2007) cannot be the result of the same agreement licensing mechanism either. In languages like Japanese and Chinese, null subjects, as well as null objects, are possible in the

absence of an agreement system. This phenomenon should therefore be analyzed in different terms, as has been done by Neeleman & Szendrői (2007). Importantly, though, there is nothing in their analysis that is incompatible with ours.

References

Arregi, Karlos. 2000. How the Spanish verb works. Handout of a talk given at LSRL 30, at the University of Florida, Gainesville.

Arregi, Karlos & Andrew Nevins. 2012. *Morphotactics: Basque auxiliaries and the structure of spellout.* Dordrecht: Springer.

Bayer, Joseph. 1984. 'COMP in Bavarian syntax.' *The Linguistic Review 3*: 209–274.

Béjar, Susana. 2003. Phi-syntax: A theory of agreement. Toronto: University of Toronto dissertation.

Benveniste, Émile. 1971. *Problems in general linguistics*. Coral Gables, FL: University of Miami Press.

Bjorkman, Bronwyn. & Hedde Zeijlstra. 2019. 'Checking up on (Φ-)Agree.' *Linguistic Inquiry 50*: 527-569.

Bobaljik, Jonathan. 2003. 'Realizing Germanic inflection: why morphology does not drive syntax.' *Journal of Comparative Germanic Linguistics 6*: 129–167.

Bobaljik, Jonathan & Höskuldur Thrainsson. 1998. 'Two heads aren't always better than one.' *Syntax 1*: 37-71.

Bohnacker, Ute. 2013. 'Null subjects in Swabian.' *Studia Linguistica 67*: 257–289.

Bonet, Eulalia. 1991. Morphology after syntax: Pronominal clitics in Romance. Cambridge, MA: MIT dissertation.

Calabrese, Andrea. 2015. Irregular Morphology and Athematic Verbs in Italo-Romance. *Isogloss. A Journal on Variation of Romance and Iberian languages*. 69-102.

Calabrese, Andrea. 2019. *Morpho-Phonological Investigations:A Theory of PF. From Syntax to Phonology in Sanskrit and Italian Verbal Systems.* Ms., University of Connecticut.

Chomsky, Noam 2000. 'Minimalist Inquiries: The Framework.' In: R. Martins, D. Michaels & J. Uriagereka (eds.) *Step by Step*. Cambridge MA: MIT Press. 89–155.

Chomsky, Noam. 2001. 'Derivation by phase.' In: M. Kenstowicz (ed.), *Ken Hale: A Life in Language*. Cambridge, MA: MIT Press. 1–53.

Cooper, Kathrin. 1994. *Topics in Zurich German Syntax*. PhD Dissertation, University of Edinburgh.

Cooper, Kathrin & Elisabeth Engdahl. 1989. 'Null Subjects in Zürich German.' *Working Papers in Scandinavian Syntax 44*: 31–44.

de Haan, Ger. 1994. 'Inflection and cliticization in Frisian.' *NOWELE 23*: 75–90.

Embick, David. 2015. *The morpheme: A theoretical introduction*. Boston and Berlin: Walter de Gruyter.

Embick, David & Rolf Noyer. 2007. 'Distributed Morphology and the Syntax-Morphology Interface.' In: G. Ramchand & C. Reiss (eds.), *The Oxford Handbook of Linguistic Interfaces*. Oxford: Oxford University Press. 289-324.

Forchheimer Paul. 1953 *The category of person in language*. Berlin: Walter de Gruyter.

Frampton, John. 2002. 'Syncretism, impoverishment and the structure of person features.' In: M. Andronis, M. Debenport, A. Pycha & K. Yoshimura (eds.), *Proceedings of CLS 38*. Chicago: Chicago Linguistic Society. 207–222.

Fuß, Eric. 2005. *The Rise of Agreement*. Amsterdam/Philadelphia: John Benjamins.

Haag-Merz, Christine. 1996. *Pronomen im Schwäbischen: Syntax und Erwerb*. Marburg: Tectum.

Halle, Morris & Alec Marantz. 1993. 'Distributed Morphology and the Pieces of Inflection.' In: K. Hale & S. Keyser (eds.), *The View from Building 20*. Cambridge MA: MIT Press. 111–176.

Harley, Heidi. & Elizabeth Ritter. 2002. 'A feature-geometric analysis of person and number.' *Language 78*: 482-526.

Heim, Irene. 2008. 'Features on bound pronouns.' In: D. Harbour, D. Adger & S. Béjar (eds.), *Phi theory: Phi-features across modules and interfaces*. Oxford: Oxford University Press. 35–56.

Holmberg, Anders. 2005. 'Is there a little pro? Evidence from Finnish.' *Linguistic Inquiry 36*: 533–564.

Holmberg, Anders. 2010. 'Null subject parameters.' In: Th. Biberauer, A. Holmberg, I. Roberts & M. Sheehan (eds.), *Parametric variation: null subjects in minimalist theory*. Cambridge: Cambridge University Press. 88–124.

Holmberg, Anders. & Ian. Roberts. 2010. 'The syntax-morphology relation.' *Lingua 130*: 111–131.

Holmberg, Anders. & Michelle. Sheehan. 2010. 'Control into finite clauses in partial null-subject languages.' In Th. Biberauer, A. Holmberg, I. Roberts & M. Sheehan (eds.), *Parametric variation: null subjects in minimalist theory*. Cambridge: Cambridge University Press. 125–152.

Huang, James. 1984. 'On the distribution and reference of empty pronouns.' *Linguistic Inquiry 15*: 531–574.

Jaeggli, Osvaldo. & Kenneth. Safir. 1989. 'The null-subject parameter and parametric theory.' In: O. Jaeggli & K. Safir (eds.), *The Null Subject Parameter*. Dordrecht: Kluwer. 1– 44.

Koeneman, Olaf. 2000. *The flexible nature of verb movement*. Doctoral dissertation. Utrecht University.

Koeneman, Olaf & Hedde Zeijlstra. 2014. 'The Rich Agreement Hypothesis Rehabilitated.' *Linguistic Inquiry 45*: 571–615.

Kratzer, Angelika. 2009. 'Making a pronoun: Fake indexicals as windows into the properties of pronouns.' *Linguistic Inquiry 40*: 187–237.

Ledgeway, Adam, 2016. 'The Dialects of Southern Italy', in Adam Ledgeway and Martin Maiden (eds.), *The Oxford Guide to the Romance Languages*. Oxford: Oxford University Press. 246-269.

Müller, Gereon. 2006. 'Pro-drop and impoverishment.' In: P. Brandt & E. Fuß (eds.) *Form, Structure, and Grammar. A Festschrift Presented to Gunther Grewendorf on Occasion of His 60th Birthday*. Berlin: Akademie Verlag. 93–115.

Neeleman, Ad. & Kriszta. Szendrői. 2007. 'Radical pro drop and the morphology of pronouns.' *Linguistic Inquiry 38*: 671–714.

Nevins, Andrew. 2003. Do person/number syncretisms refer to negative values? Paper presented at the annual meeting of the Linguistic Society of America, Atlanta, GA, January 2–5.

Nevins, Andrew. 2007. 'The representation of third person and its consequences for person-case effects.' *Natural Language & Linguistic Theory 25*: 273–313.

Newmeyer, Frederick J. 2004. Against a parameter-setting approach to language variation. In Language Variation Yearbook, Volume 4 , ed. by Pierre Pica, Johan Rooryck and J. van Craenenbroek, 181-234. Amsterdam: Benjamins.

Oltra-Massuet, Isabel. 1999. 'On the Constituent Structure of Catalan Verbs,' in K. Arregi, B. Bruening, C. Krause, and V. Lin, eds., Papers in Morphology and Syntax, Cycle One, volume 33 of MIT Working papers in Linguistics, MITWPL, Cambridge, MA, 279–322.

Pesetsky, David & Esther. Torrego. 2007. 'The syntax of valuation and the interpretability of features.' In: S. Karimi, V. Samiian & W. Wilkins (eds.), *Phrasal and Clausal Architecture: Syntactic derivation and interpretation*. Amsterdam: Benjamins. 262–294.

Pinker, Steven. 1984. *Language learnability and language development*. Harvard: Harvard University Press.

Preminger, Omer. 2014. *Agreement and its failures*. Cambridge, MA: The MIT Press.

Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.

Roberts, Ian. 2010. 'Varieties of French and the null subject parameter.' In: Th. Biberauer, A. Holmberg, I. Roberts & M. Sheehan (eds.), *Parametric variation: null subjects in minimalist theory*. Cambridge: Cambridge University Press, 1–57.

Rosenkvist, Henrik. 2009. 'Referential Null Subjects in Germanic Languages – an Overview.' *Working Papers in Scandinavian Syntax 84*: 151–180.

Saab, Andrés. 2008. *Hacía una teoría de la identidad parcial en la elipsis*. Doctoral dissertation, University of Buenos Aires.

Sauerland, Uli. 2002. 'The present tense is semantically vacuous.' *Snippets 6*: 12–13.

Sauerland, Uli. 2008. On the semantic markedness of phi-features. In: D. Harbour, D. Adger & S. Béjar (eds.), *Phi theory: Phi-features across modules and interfaces*. Oxford: Oxford University Press. 57–82.

Sheehan, Michelle. 2016. Subjects, null subjects and expletives in Romance. In S. Fischer & S. Gabriel (eds.), *Manuals of Romance Linguistics (MRL): Grammatical Interfaces*, 329-362.

Taraldsen, Knut Tarald. 1978. *On the NIC, vacuous application and the That-Trace Filter*. Ms., MIT.

Trutkowski Ewa. 2016. *Topic drop and null subjects in German*. Berlin: De Gruyter.

Van Alem, Astrid. 2020. 'Complementizer agreement is not allomorphy: A reply to Weisser (2019).' *Glossa: A Journal of General* Linguistics 5: 44.

Van Alem, Astrid. forthcoming. *Life of phi*. PhD Dissertation, Utrecht University.

Vainikka, Anne & Yonata Levy. 1999. 'Empty subjects in Hebrew and Finnish.' *Natural Language and Linguistic Theory 17*: 613–671.

Von Stechow, Arnim (2003). Feature deletion under semantic binding: Tense, person, and mood under verbal quantifiers. In M. Kadawaki & S. Kawahara (eds.), *Proceedings of NELS* 33, 379–404. Charleston: BookSurge Publishing.

Tamburelli, Marco. 2006. Remarks on Richness. *UCL Working Paper Linguistics*, 18.

Weiß, Helmut. 2005. 'Inflected complementizers in Continental West Germanic dialects.' *Zeitschrift für Dialektologie und Linguistik 72*: 148–166.

Wratil, Melani. 2011. Uncovered pro - On the development and identification of null subjects. In Wratil, M. & P. Gallmann (eds.) *Null pronouns*. Berlin/Bosten: De Gruyter. 99-140.