Language and Communication in the Digital Age

Language and Communication in the Digital Age

Edited by

Marinela Burada, Oana Tatu
and Raluca Sinu

Language and Communication in the Digital Age

Edited by Marinela Burada, Oana Tatu and Raluca Sinu

This book first published 2023

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

TABLE OF CONTENTS

Chapter III. Approaches to Language Teaching and Acquisition

CHAPTER I:

PERSPECTIVES ON LANGUAGE STRUCTURE

PERSPECTIVES ON LANGUAGE STRUCTURE: INTRODUCTION

RALUCA SINU

This chapter investigates different aspects pertaining to the structure of natural languages. The contributions gathered here deal with language description at the phonological, lexical, morphological and syntactic levels, approaching topics such as the agreement between subject and predicate, doxatic predicates, colour degree achievement verbs, or comment clause parentheticals. Many of the authors resort to contrastive analyses to highlight common points and differences between Romanian, on the one hand, and other European languages, on the other, or Gulf Pidgin Arabic and its relationship to Gulf Arabic.

A New Perspective on Pro Drop in Romanian (Olaf Koeneman and Hedde Zeijlstra) starts from the observation that a language with rich agreement morphology can leave out a subject in a tensed clause, for example, Italian, while in English this results in cases of ungrammaticality. Within the framework of a new theory on pro drop and relying on examples from Icelandic and Romanian, the authors set out to prove that pro drop is only possible if a language expresses tense and agreement by separate morphemes. The detailed overview of morphemic expressions of tense and agreement in the article shows that languages like Italian, Spanish and Romanian mark tense and agreement in a transparent way, which is not true for English and Icelandic. This is followed by a case study of Romanian focusing on the agreement forms in the present and imperfect tense of the first and the fourth conjugations. The authors resort to impoverishment rules to explain the agreement alternations between different parts of the paradigm, stating that these rules apply to Romanian, but not to Icelandic or to English. They also zoom in on the problem of the 3rd person singular form in the present tense context which does not return in the 3rd person singular in the imperfect context in Romanian, accounting for it in two different ways, i.e. the deletion account and the null 3SG account. However, the overall aim of the article is to understand Romanian in contrast to Icelandic, to explain why, despite the similarities between the structures of the two languages, Romanian has pro drop, while

Icelandic has not. This type of contrastive theory is considered by the authors the only way to increase the predictive power of pro drop theory.

On Gender Mismatches in Superlative and Quantified Partitives (Mihaela Tănase-Dogaru) approaches the field of partitivity in Romanian. The paper begins with a discussion of gender mismatches in French (cf. Sleeman and Ihsane 2016) and German (cf. Westveer 2021) for two categories of animate nouns, i.e. the class containing alternating forms - the *professor* type -, and the class containing epicene forms - the *dean* type. Then the author describes the situation for Romanian, attempting to demonstrate that these types of gender mismatches are possible both in superlative and quantified partitive constructions, and seem to involve the presence of a syntactic projection for Gender in both determiner phrases.

In **Romanian Doxastic Predicates: Negation, Mood Shift and Types of Belief** (Maria Aurelia Cotfas), the author discusses the mood selection in the case of doxastic predicates in Romanian. The starting point of the paper is the veridicality/non-veridicality framework built by Giannakidou and Mari (2021), who distinguish between two types of commitment: solipsistic and conjectural. The aim of the research is to prove that there is mood shift in Romanian as well as to signal the difference in commitment: the indicative yields (subjective) veridicality, while the subjunctive signals conjecture. In support of this idea, both theoretical and empirical evidence is provided. The latter comes under the form of small-scale questionnaires in which Romanian respondents were asked to assess the various readings of several sentences. The article concludes, among others, that in Romanian the speakers can choose between the indicative mood in the case of affirmative and negated doxastic verbs, and the subjunctive in the case of negated matrix attitude verbs.

Moving on, **Colour Degree Achievement Verbs on a Quantitative Reading Are Closed Scale Degree Verbs: Romanian Native-Speaker Data** (Ioana Stoicescu) presents novel data on the semantic behavior of a subtype of change of state predicates, specifically degree verbs of colour, which have not been previously investigated in Romanian. The objective is to determine whether colour degree verbs on a quantitative reading display a similar behaviour to three other types of scalar verbs, namely absolute minimum standard verbs, absolute maximum standard degree achievements, and open class degree achievements on various entailment and modification patterns. The analysis builds on previous work on adjectival derivation and scalar structure, with a focus on closed scales. The findings indicate that Romanian colour degree verbs are associated with closed scales on a quantitative reading, exhibiting similar patterns to

absolute minimum standard degree achievements, while also showing a preference for the telic interpretation of absolute maximum standard verbs. The author concludes that further research is needed, with a larger pool of respondents, in order to investigate the entailment and modification patterns of colour verbs on a qualitative reading.

The Semantic and Syntactic Representation of Comment Clause Parentheticals in Romanian (Andreea Codrina Tănase) represents a contribution to the ongoing debate on the syntactic status of parenthetical constructions, and brings evidence for the integrated approach, which posits that parentheticals are syntactically integrated into the host sentence despite their status as independent units. The study begins with a classification of comment clause parentheticals in modern Romanian, and goes on to discuss existing approaches in the literature on parentheticals. The author argues that the integrated approach is more appropriate, and provides examples in Romanian that demonstrate how syntax can reflect meaning shift. The integrated approach is further supported by the discussion on the ambiguities that arise when an anaphoric pronoun is faced with choosing between antecedents that are sisters in the tree representation. With a view to showing how parentheticals act as meaning modifiers in non-assertive host sentences, the author also includes a brief corpora study. All things considered, the paper argues that the syntactic and semantic features of parentheticals indicate that they are integrated into the host sentence, and that syntax should always reflect any changes in language.

Concentrating on the status of a pidgin language, **Gulf Pidgin Arabic: Transient Learner Variety of Gulf Arabic or Pidgin?** (Andrei A. Avram) provides an overview of a number of issues which should be taken into consideration in deciding whether Gulf Pidgin Arabic is a learner variety of Gulf Arabic. The paper looks at, among others, the phonetic realizations of some of the marked consonants of Gulf Arabic; word order; the forms of pronouns, adjectives and verbs; and several grammaticalization phenomena. Based on the data collected from speakers in previous studies, the correlation between the length of stay in the countries of the Arab Gulf and a better approximation of Gulf Arabic, presumed to be the target language, was rather weak, which prompts the author to conclude that Gulf Pidgin Arabic is not a learner variety of Gulf Arabic, but rather a variety in its own right, although it has not yet undergone stabilization in its development.

A New Perspective on Pro Drop in Romanian

Olaf Koeneman and Hedde Zeijlstra

Abstract. A strong correlation can be observed between the availability of rich agreement morphology in a language and the ability to leave out the subject from a tensed clause (cf. Taraldsen 1978, Rizzi 1982, and many others), but it has proven hard to understand this correlation in its details. In this paper, we will provide a new theory on pro drop, which holds that pro drop is only possible if a language expresses tense and agreement by separate morphemes (Koeneman and Zeijlstra, under review). We will consequently apply this theory to Icelandic and Romanian. Although these languages share important properties on the surface (they look equally rich), only the latter licenses argumental null subject. We will show how our proposal can correctly derive this difference.

1 Introduction

It is a well-known fact that a language like Italian can leave out a subject in a tensed clause (i.e., it licenses null subjects, or pro drop), whereas this leads to ungrammaticality in English. The contrast is shown in (1):

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

This difference can be linked to the fact that Italian has a rich agreement system whereas English does not. As can be seen in (2), Italian shows six agreement distinctions in the present tense paradigm, whereas English only shows two:

(2)   Present and imperfect/past tenses in Italian and English

	ITALIAN		ENGLISH	
	amare ('to love)		to talk	
	present	*imperfect*	*present*	*past*
1SG	amo	amavo	talk	talked
2SG	ami	amavi	talk	talked
3SG	ama	amava	talks	talked
1PL	amiamo	amavamo	talk	talked
2PL	amate	amavate	talk	talked
3PL	amano	amavano	talk	talked

The link between rich agreement and null subjects makes a lot of intuitive sense: if the agreement form expresses (almost) the same information as the subject does, the latter can easily go unexpressed. There is, however, a fundamental issue with this idea. The 3rd person singular form in English is as unique to the English agreement system as the 3rd person singular form in Italian is. This raises the question why English does not allow a null in at least 3SG contexts.

The usual solution is to formulate a binary parameter and assume that a negative setting (a consequence of poor agreement under some definition) disallows null subjects across the board, whereas a positive setting (a consequence of rich agreement under some definition) will allow the language to generate null subjects across the board. The problem for such a parametric approach, however, is the existence of so-called partial pro drop languages, which allow null subjects in some but not all contexts. Frisian, for instance, allows null subjects in 2SG contexts but not elsewhere (cf. De Haan 1984). Bavarian dialects are similar. Although some varieties also allow null subjects in 1PL and 2PL contexts, no variety allows them across the board (cf. Bayer 1984). Partial pro drop languages therefore show that licensing null subjects does not have to be an all-or-nothing affair and can be determined per (person/number) context. If so, why does English not allow it in 3SG contexts?

This problem aggravates when one considers a richer non-pro drop language like Icelandic, which has five distinctions in its agreement paradigm (cf. (3)). Despite the fact that it consequently marks four contexts uniquely, null argumental subjects are not licensed. This is especially noteworthy in the face of Romanian, which also shows five distinctions (shown for the first conjugation in (3)).

(3)   Present and past/imperfect tenses in Icelandic and Romanian

	ICELANDIC		ROMANIAN	
	heyra ('to hear')		cânta ('to sing')	
	present	*past*	*present*	*past (imperfect)*
1SG	heyr-i	heyrði	cânt	cântam
2SG	**heyr-ir**	heyrðir	cânți	cântai
3SG	**heyr-ir**	heyrði	**cântă**	cânta
1PL	heyr-jum	heyrðum	cântăm	cântam
2PL	heyr-ið	heyrðuð	cântați	cântați
3PL	heyr-a	heyrðu	**cântă**	cântau

In contrast to Icelandic, Romanian is a full pro drop language like Italian. Now, it can be observed that the two contexts that in Romanian are syncretic in the present tense (3SG-3PL) are no longer syncretic in the imperfect tense, so one could perhaps argue that this is important for the licensing of null subjects. Note, however, that the exact same is true for Icelandic (cf. (3)). Given these similarities, it is hard to understand why Romanian is a full-fledged pro drop language, whereas Icelandic does not allow argumental null subjects at all.

In the next section, we will present a new proposal on how the link between agreement and pro drop is to be understood. We will argue that a language like Italian expresses tense and agreement in a transparent way, whereas this is not the case in languages like English and Icelandic. With this proposal in place, section 3 will look at Romanian in detail and show that it patterns with Italian and Spanish rather than with English and Icelandic. Section 4 deals with a central issue that Romanian poses for our proposal, having to do with an agreement alternation in the 3rd person singular. Two consequences of our analysis for understanding pro drop in Romanian are discussed in section 5, and section 6 concludes the analysis.

2 Mono- versus bi-morphemic expression of tense and agreement

We generally agree with the literature stating that rich agreement is a prerequisite for the licensing of pro drop.[1] Given the existence of the partial pro drop phenomenon, we assume that null subjects are in principle licensed contextually: if the agreement information on the verb is identical to the information expressed by a pronominal subject, the subject can go unexpressed.[2] These two assumptions together clearly overgenerate. They

predict that any context that is uniquely marked (e.g. 3SG contexts in English and Icelandic, as well as all plural contexts in Icelandic) license null subjects, contrary to fact. There must therefore be another prerequisite for such licensing, which apparently Italian, Spanish and Romanian meet but English and Icelandic do not. This prerequisite, we propose, is the transparent marking of tense and agreement. What we mean by this is the following.

If you look at the present and imperfect tense systems of Italian and Spanish, we can clearly identify the past tense markers as -v- and -b- respectively. The agreement forms we discern in the present tense follow this past tense marker in the imperfect. This means that the forms at the surface provide strong evidence for the existence of two separate morphemes underlying these forms, one expressing tense features and another expressing agreement features. This is in contrast to English, where we observe that the only visible agreement form, the 3SG -s, does not return in the past tense.

(4)   Present and past/imperfect tenses in Italian, Spanish and English

	ITALIAN		SPANISH		ENGLISH	
	present	*imperfect*	*present*	*imperfect*	*present*	*past*
1SG	amo	amavo	amo	amaba	talk	talked
2SG	ami	amavi	amas	amabas	talk	talked
3SG	ama	amava	ama	amaba	talks	talked
1PL	amiamo	amavamo	amamos	amábamos	talk	talked
2PL	amate	amavate	amáis	amabais	talk	talked
3PL	amano	amavano	aman	amaban	talk	talked

This has led Bobaljik and Thráinsson (1998) to conclude that English does not have separate morphemes for tense and agreement but a single one expressing both properties, I(NFL). In such an analysis -s and -ed are in direct competition for insertion into this one morphemic slot. The difference between Italian and English would then boil down to different settings of the so-called Split-IP Parameter: whereas English syntax generates INFL, Italian generates separate slots for tense and agreement. This is illustrated in (5) below.

Given this background, we can now hypothesize that the difference between a bi-morphemic and a mono-morphemic expression of tense and agreement also underlies the difference between having or not having pro drop (see Saab 2008, for a similar idea). More specifically, we can say that English does not license null subjects because, even in the context in which the agreement form is a unique marker (i.e., 3SG contexts), the

morpheme expressing the relevant agreement features also expresses tense features. This morpheme is therefore overspecified, as it crucially expresses features (namely tense features) that are semantically incompatible with the subject that it is supposed to license. For null subjects to be possible, then, a split IP and a separate agreement morpheme constitute the relevant prerequisite.

(5)   a. ITALIAN SPLIT-IP   b.   ENGLISH UNSPLIT-IP

```
        AgrP                                      IP
       /    \                                    /  \
  Subject   Agr'                           Subject   I'
           /    \                                   /  \
        Agr      TP                               I    VP
         ¦      /  \                              ¦
         ¦    ..    T'                            ¦
         ¦        /   \                           ¦
         ¦       T     VP                         ¦
         ¦       ¦                                ¦
        -o      -v                             -s/-ed
```
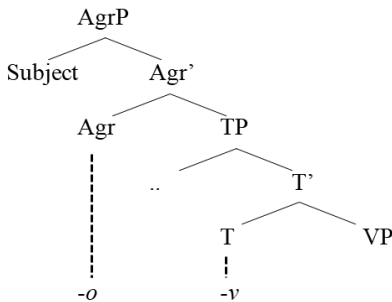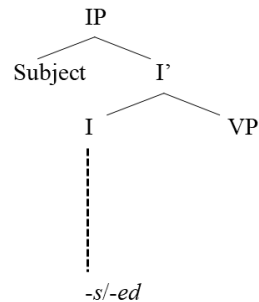
When we now turn to Icelandic, it is not immediately obvious how the analysis makes the right prediction. Take a look at (3) again. We could identify -ð- as the past tense marker that is followed by agreement forms, and these forms at least partially resemble forms we also see in the present tense. In short, Icelandic at first view looks like a split-IP language, and this is in fact how Bobaljik and Thráinsson analyze it. Of course, given our hypothesis, this would erroneously predict that Icelandic licenses argumental null subjects, contrary to fact.

However, there is one property that Icelandic shares with English: the form that appears in the 3SG present tense does not return in the 3SG past tense.[3] English has -s in the present tense but no overt 3SG agreement marker in the past. Icelandic has -ir as the 3SG marker in the present tense, but -i in the past tense. In other words, there is a morpho-phonological contrast in these 3SG contexts and this is in contrast with for instance Italian and Spanish. As can be observed in (4), the form in the 3SG context in the present tense is *(am)a*. The *-a* following the stem can be treated as the theme vowel that we also see in the infinitive (*am**a**re* and *am**a**r*) and in the plural contexts (1PL *ami**a**no/am**a**mos*, 2PL *am**a**te/am**á**is*, 3PL *am**a**no/am**a**n*).[4] Therefore, the 3SG agreement marker is null, contrasting with -o and -i/-as in the 1SG and 2SG, respectively. Crucially, this null form returns in the imperfect tense, where we see the exact same contrasts

(we will return to the unexpected *-a* in 1SG imperfect contexts in Spanish in section 3). In contrast to Icelandic and English, then, Italian and Spanish show no difference between the forms occurring in 3SG present and imperfect contexts. This particular property that sets Germanic agreement systems apart makes it harder to maintain a transparent analysis for tense and agreement, as we will now see.[5]

Let us look in more detail at the rules that insert concrete forms into the morphemic slots. We will make the standard assumption that the third person is the non-person in the sense of Forchheimer (1953) and Benveniste (1971) (see also Harley and Ritter 2002 and Preminger 2014). This entails that the form appearing in the 3SG context will function as the underspecified elsewhere form of the paradigm, expressing no feature values. Now, let us aim for a bi-morphemic analysis of Icelandic, starting with the following insertion rules for the singular part of the Icelandic paradigm:

(6)   *-i* <>   [Agr: speaker]
           -ir   <>   [Agr:]

What the grammar of Icelandic now needs to capture is the following two properties: (i) 3SG *-ir* becomes *-i* in the past tense; (ii) *-ir* remains *-ir* in the 2SG past context. We first postulate a separate entry for the past tense marker *-ð*.

(7)   *-ð* <>   [T: past]

Property (i) can then be achieved by postulating a context-sensitive rule which spells out the same features as *-ir* does in the present tense but inserts a different form. Such rules are needed generally to account for alternations between paradigms. One example is the 1PL context in Italian (cf. (4)), where the 1PL marker is *-iamo* in the present tense and *-amo* in the imperfect. This could be captured by adopting the following two rules[6]:

(8)   *-iamo*   <>   [Agr: speaker, plural]
           -amo   <>   [Agr: speaker, plural] / [T: past]

Along the same lines, we can now postulate a context-sensitive rule for Icelandic to capture the fact that 3SG *-ir* in the present tense corresponds to *-i* in the past tense:

(9)   *-i* <>   [Agr:] / [T: past]

Note, however, that under this analysis the *-i* appearing in 3SG past tense contexts is different from the *-i* appearing in 1SG past tense contexts: (9) must be assumed alongside the spell-out rule referring to [speaker] in (6). Therefore, the analysis postulates that these two instances of *-i* are accidentally homonymous instead of syncretic: it fails to capture their similarity directly. At the same time, the analysis now predicts that *-ir* becomes *-i* in 2SG past contexts too. In other words, it fails to capture property (ii). In order to prevent (9) from generating *-i* in 2SG past tense contexts, we have to assume that the *-ir* forms appearing in 2SG and 3SG present tense contexts are homonymous too and that the rule in (9) only competes with the one in (10)b.

(10)   a. *-ir*   <>   [Agr: addressee]
        b.   *-ir*   <>   [Agr:]

We conclude that a transparent, bi-morphemic analysis of tense and agreement can be maintained for Icelandic but at the cost of creating two homonym pairs.

Similar but different issues arise for English. In order to account for the fact that the 3SG *-s* does not return in the past tense, it must be assumed that *-s* competes with a null allomorph functioning as the elsewhere form in the past tense.

(11)   -∅   <>   [Agr:] / [T: past]

This null allomorph will in addition be homonymous with the null form appearing in the 1SG and 2SG present and past contexts:

(12)   -∅   <>   [Agr: participant]

We conclude that a bi-morphemic analysis can be maintained for English too, but at the cost of postulating a null allomorph and a homonymous pair.

There is an alternative analysis accounting for the absence of the 3SG present tense form in the past, namely one that assumes that the 3SG *-ir* form in Icelandic does not return in the 3SG past context because it directly competes with *-ði*. This analysis treats *-ði* as one form rather than two, and requires the existence of an underlying morpheme expressing both tense and agreement features. In other words, the analysis treats Icelandic as a non-split-IP language. In English, *-s* would directly compete with *-ed* and this boils down to the analysis proposed by Bobaljik and Thráinsson (1998). Note that such an analysis does not have to postulate

(11) because *-s* is specified for present tense (cf. (13)). In addition, the mono-morphemic analysis also does not need to postulate a separate null form that spells out present tense (because the forms -Ø and -s in (13) already do that), thereby economizing on two null forms in total.[7] The overall advantage of the mono-morphemic analyses is that they do not postulate unwanted null forms and homonyms. The singular parts of the paradigms can be captured by the following rules.

(13)

Icelandic			English		
-i	<>	[I: present, speaker]	-Ø	<>	[I: present, participant]
-ir	<>	[I: present, __]	*-s*	<>	[I: present, __]
-ði	<>	[I: past, __]	*-ed*	<>	[I: past, __]
-ðir	<>	[I: past, addressee]			

In sum, we have briefly explored the bi-morphemic and mono-morphemic takes on Icelandic and English, and shown that in principle both are possible. From a theoretical perspective, it may be hard, or random, to choose one analysis over the other. The mono-morphemic analyses may economize on null allomorphs and homonyms, but only the bi-morphemic analyses directly capture the fact that all past forms contain *-ð* and *-ed*. After all, these forms are not listed as separate [T: past]-markers in the mono-morphemic analyses. If we reason from the perspective of child language acquisition, however, it becomes clearer why the mono-morphemic analyses have a plausible appeal. Take English, where the child observes that *-s* and *-ed* are in complementary distribution. To capture this, the child can either assume direct competition between the two overt forms that (s)he has direct evidence for (namely *-s* and *-ed*) and settle for a mono-morphemic analysis, or create a competition effect between *-s* and an invisible form, the null allomorph in (11), for which there is no direct evidence. For the language learning child, the mono-morphemic analysis is therefore the more obvious choice as it relies on direct evidence. In the same vein, the mono-morphemic analysis is a way for the Icelandic child to not have to postulate superfluous homonymous pairs.

We can summarize our proposal in the following way:

(14)   If forms x and y are in complementary distribution, an analysis that takes x and y to be in direct competition takes precedence over an analysis that accounts for the effect by postulating a form z, where z is either a null form or a homonym.

It is important to understand that this proposal does not put any ban on null forms or homonym in general. They can be part of the child's analysis but are dispreferred if they are postulated as a way to capture a complementary distribution effect *between two other forms*. In that event, there is always an alternative analysis available, namely one that assumes direct competition between the two forms. Any null form or homonym that is not postulated to capture a complementary distribution effect of two other forms is therefore allowed, as (14) has nothing to say about these. To make this concrete, let us look at an unproblematic null form and homonym, and contrast them with the problematic ones.

As an example of an unproblematic null form, take for instance the Italian paradigm in (4). Italian is like any other language discussed here in not expressing the present tense with an overt marker. Postulation of a non-overt one is straightforward, as such a null form stands in direct contrast with the *-v* that marks the imperfect. This null marker for the present tense is paradigmatically licensed and adopted as the spell out of present tense, which no other form in the paradigm would otherwise take care of. The status of this null form is therefore fundamentally different from the one that is made problematic by (14). The competing x and y forms in English are *-s* and *-ed*, respectively, and capturing their complementary distribution by postulating an additional null form is dispreferred, given (14), and therefore problematic. In Italian, however, the x and y forms in competition are the present tense marker -Ø and the imperfect marker *-v*. The statement in (14) does not even bear on this null form, as it simply does not rule out that either x or y itself is a null form.

A similar point can be made for homonymy. If there are two identical forms appearing in different slots of the paradigm, these forms will be analyzed as syncretic if the contexts in which we observe them shares features. Take as an example the Icelandic past tense:

(15)   Icelandic past tense and German present tense

	ICELANDIC		GERMAN	
	heyra ('to hear')		spielen ('to play')	
	past		*present*	
	SG	PL	SG	PL
1	heyrði	heyrðum	spiele	spielen
2	heyrðir	heyrðuð	spielst	spielt
3	heyrði	heyrðu	spielt	spielen

In Icelandic, we can observe that the 1SG and 3SG forms are identical. Since these appear in the same column, they obviously share a feature, namely the number feature [singular]. It is possible, therefore, to treat these two forms as syncretic. An analysis that does not do so and postulates two *-i* (or *-ði*) forms instead assumes the existence of a homonymous pair that can in principle be avoided. Now, we saw in the analysis of Icelandic that there is a reason for postulating two *-i* forms, namely to capture the absence of the 3SG present tense form *-ir* in the 3SG past context under a bi-morphemic analysis. Since this essentially involves a complementary distribution effect of two forms, *-ir* and *-ði*, (14) applies and the analysis is dispreferred. This Icelandic example stands in contrast to the German example in (15). Note for this language that the 3SG and 2PL contexts share the same form, namely *-t*, but these contexts are not in the same column, nor in the same row. In other words, they do not share a person or number feature. Therefore, postulating a homonymous pair of *-t* forms is unproblematic for lack of a better analysis.[8] Again, the statement in (14) does not even bear on this case, as the homonymous *-t* pair is not postulated so as to capture a complementary distribution effect of two other forms.

This distinction between avoidable and non-avoidable null forms and homonyms becomes important when we turn to Romanian.

To conclude, Icelandic and English can be analyzed bi-morphemically and mono-morphemically but from the view of a language learning child there are clear reasons to prefer mono-morphemic ones for these languages, which we have formulated in (14). We can now turn to the analysis of Romanian. Recall that this language shares with Icelandic the number of agreement contrasts in the present tense, and the fact that contexts that are syncretic in the present tense are no longer so in the past tense (cf. (3)). Since Romanian is a full-fledged pro drop language and Icelandic is not, the question is whether a bi-morphemic analysis can be maintained for Romanian. This will be explored in the next section.

3 A morphological analysis of Romanian

It is beyond the scope of this paper to provide an exhaustive analysis of the Romanian inflectional system. Below, we will focus on the agreement forms in the present and imperfect tense of two different conjugations, the first and the fourth one. These show a different distribution of distinct forms, representative of the patterns found in Romanian at large. They constitute paradigms which are fully productive and will therefore provide the child with the necessary information to

choose between a split- or unsplit-IP analysis for the language. Wherever relevant, we will refer to other conjugations.

Let us start by looking at the two conjugations together.[9] This leads to two important observations. The first one is that we can identify -*a* as the imperfect marker contrasting with no overt tense marker for the present tense. Since -*a* occurs in all contexts of the imperfect and irrespective of conjugation, it cannot be analyzed as a conjugation-specific theme vowel. The second important observation is that different conjugations have different agreement systems. For the first conjugation, it can be observed that each slot in the paradigm has its own agreement marker with the exception of the 3SG and 3PL forms, which are the same. If we look at the imperfect, however, we first of all see that the forms for 3SG and 3PL are distinct again (-*a* versus -*au*). In addition, we see that the 1SG and 1PL forms now look identical, in contrast with what happens in the present tense. In the fourth conjugation, we observe a different pattern. In the present tense, the 3PL form is not identical to the 3SG form but to the 1SG form. In the imperfect tense, however, the 1SG and 3PL contexts show different agreement forms (-*am* versus -*au*), whereas the 1SG and 1PL are similar again, just like in the first conjugation.

(16) Romanian present and imperfect agreement paradigms

	FIRST CONJUGATION		FOURTH CONJUGATION	
	a cânta ('to sing')		a fugi ('to run')	
	present	*past (imperfect)*	*present*	*past (imperfect)*
1SG	cânt	cântam	fug	fugeam
2SG	cânți	cântai	fugi	fugeai
3SG	cântă	cânta	fuge	fugea
1PL	cântăm	cântam	fugim	fugeam
2PL	cântați	cântați	fugiți	fugeați
3PL	cântă	cântau	fug	fugeau

The question is whether we can provide an analysis for Romanian that captures these agreement alternations between present and imperfect tense in a bi-morphemic system (i.e., with all insertion rules either referring to tense or agreement features but not to a mixture of these two) while at the same time refrain from postulating avoidable null forms and homonym pairs, as these properties can trigger a switch to a mono-morphemic system, as we have hypothesized for Icelandic and English. We will first provide a general analysis of the first and fourth conjugations. These analyses will treat Romanian as a bi-morphemic language and will capture the basic facts. After that, we will focus on one

problem for this analysis: as can be observed in (16), the 3SG form (i.e., *-ă* in the first and *-e* in the fourth conjugation) does not seem to return in the 3SG imperfect context. This initially suggests that Romanian runs into the same trap as English and Icelandic, raising the question why Romanian is not a mono-morphemic language without argumental pro drop. We will then sketch the two logical solutions available to us and show that the most plausible analysis is one in which the 3SG markers are in fact generated in imperfect contexts but targeted by phonological deletion, a process that can be independently justified for Romanian.

Let us start with an analysis of the first conjugation. Under the assumption that *-a* is the imperfect marker, the spell-out rules for tense are straightforward as seen in 17 below.

(17)   *Tense*
          -∅   <>   [T:] (present)
          -*a*   <>   [T: past]

Like for any other language discussed here, the null form spelling out present tense is unproblematic. There is no other form spelling out this feature (it is not a null allomorph) and the null form stands in direct contrast with the *-a* marking the imperfect.

The spell-out rules for the agreement forms are less straightforward and require more discussion. We propose the following spell-out rules:[10]

(18)

Agreement			*Tense*		
-∅	<>	[Agr: speaker, singular]	-∅	<>	[T:] (present)
-*i*	<>	[Agr: addressee]	-*a*	<>	[T: past$_{\text{IMPERFECT}}$]
-*ă*	<>	[Agr:]			
-*m*	<>	[Agr: speaker]	*Theme vowels*		
-*ţi*	<>	[Agr: addressee, plural]	-*a*	<>	[Th] / Conjugation 1
-*u*	<>	[Agr: plural]/[T: past$_{\text{IMPERFECT}}$]	-*ă*	<>	[Th] / Conjugation 1, _[m]

In order to account for the same form appearing in the 3SG and 3PL present tense contexts, we treat this similarity as syncretic by taking the form inserted in these contexts as the elsewhere form, which expresses no feature values. In the other slots, specific forms appear, and this is captured by the unique feature values that each form spells out: *-i* spells out [Agr: addressee] and *-ţi* [Agr: addressee, plural], etc. The theme vowel of the first conjugation is *-a* but in the present tense only occurs in 1PL and 2PL contexts.[11] Allomorphic theme vowel *-ă* occurs in 1PL contexts, presumably related to the phonological properties of the nasal agreement ending *-m*.

If we now turn to the imperfect tense, an issue appears: the 1SG and 1PL contexts show the same form, in contrast to what we observe in the present tense. In order to account for the fact that 1SG and 1PL contexts share the same form in the imperfect but not the present tense, we postulate *-m* as the form that spells out [speaker] but not a number feature. That allows us to treat the form in these two contexts as syncretic. However, we must also account for the fact that in the present tense 1SG contexts do not have *-m* but a null form. We therefore postulate a form *-Ø* that spells out [speaker, singular]. Since the features underlying this form are more specific than the ones underlying *-m*, it blocks the latter from being inserted. In order to account for the fact that *-m* replaces *-Ø* in the 1SG imperfect context, we propose that the [singular] feature is impoverished in the past tense by the impoverishment rule in (19).

(19)   [Agr: speaker, singular] → [Agr: speaker] / [T: past]

This blocks *-Ø* from being inserted and *-m* is inserted instead as the less specified form compatible with that context.

Like context-sensitive spell-out rules, impoverishment rules (cf. Bonet 1991) are devices that can be used to account for agreement alternations between different parts of the paradigm. Specific for impoverishment rules is that they apply in marked parts of the paradigm (such as the plural or past tense dimension, cf. Aalberse and Don 2007), and lead to the appearance of unmarked forms over marked forms in those contexts. In other words, they capture the fact that another form, already part of the paradigm, fills the place of the disappearing form. Another example from Romance would be the Spanish imperfect, where we see that the contrast between 1SG and 3SG in the present tense disappears in the imperfect:

(20)   Spanish agreement in present and imperfect tense

		SPANISH	
	amar ('to love')		
	present	*imperfect*
1SG	amo	amaba
2SG	amas	amabas
3SG	ama	amaba

We can account for this pattern in a similar way as we dealt with the pattern in Romanian: by impoverishment of a feature that is linked to the more specific form, in this case *-o*. The partial grammar (which treats all /a/-vowels following the stem *am-* as theme vowels, cf. Oltra-Massuet 1999) would look as follows:

(21) *Agreement*   *Impoverishment*

    -o   <>   [Agr: speaker]   [Agr: speaker] ➔ [Agr:]/[T: past]

    -s   <>   [Agr: addressee]

    -∅   <>   [Agr:]

With impoverishment rules as a device to account for agreement alternations, an obvious question arises: what blocks us from using it to account for the disappearance of the present tense 3SG forms in the past tense in languages like English and Icelandic? This could perhaps allow us to uphold a bi-morphemic analysis for these languages without having to postulate avoidable null forms and homonyms. This would be an unwanted possibility, as the consequence would be that English and Icelandic are predicted to allow pro drop in at least those contexts that have a unique agreement marker.

The reason for the unavailability of impoverishment for English and Icelandic is as follows. If we look at Romanian and Spanish, we see that the alternations take place in 1SG contexts. These are contexts that express the [speaker] feature and in Romanian also the [singular] feature. These features can subsequently be targeted by impoverishment, and the consequence is that a less marked form (the 1st person form in Romanian, and the 3SG elsewhere form in Spanish) will appear in that slot instead. When we look at Icelandic and English, however, we see that the alternation takes place in the 3SG context. This means that impoverishment should target any feature that is linked to this context. However, it is precisely the 3SG context which is the unmarked context expressing no feature values.[12] Hence, there is nothing that impoverishment could target. This means that impoverishment rules are a useless device to account for agreement alternations happening in 3SG contexts and the only remaining device is a context-sensitive rule. Such a rule must then postulate a null allomorph for English and a homonymous form for Icelandic, and this raises the acquisitional conundrum that we discussed.

Another property of the proposed analysis that is worth pointing out is the fact that in the present tense there is no dedicated form for the 3PL context. This is why the elsewhere form *-ă* also surfaces there. In the imperfect, however, there is a dedicated form, *-u*, which must be inserted

by a context-sensitive rule. By restricting insertion of -*u* to imperfect contexts, we ensure that it is not inserted in 3PL present tense contexts.

Let us now turn to the fourth conjugation. Many aspects of our analysis of the first conjugation carry over to our analysis of the fourth one, but there are some specific details we have to focus on. We propose the following spell-out rules:

(22)

Agreement			Tense		
-∅	<>	[Agr: speaker, singular]	-∅	<>	[T:] (present)
-*i*	<>	[Agr: addressee]	-*a*	<>	[T: past$_{IMPERFECT}$]
-*e*	<>	[Agr:]			
-*m*	<>	[Agr: speaker]	*Theme vowel*		
-*ți*	<>	[Agr: addressee, plural]	-*i*	<>	[Th] / Conjugation 4
-∅	<>	[Agr: plural] / Conjugation 4			
-*u*	<>	[Agr: plural] / [T: past$_{IMPERFECT}$]			

The main difference between the agreement systems of the first and fourth conjugations is the distribution of similar forms. Whereas in the first conjugation identical forms occur in 3SG and 3PL contexts, in the fourth conjugation identical forms appear in 1SG and 3PL contexts, which both display a null form. Since these two contexts do not form a natural class, the only way to analyze the null form as syncretic is by taking it to be the elsewhere form. We have assumed throughout, however, that the form occurring in the 3SG context always functions as the elsewhere, and this assumption was part of the explanation for why languages like English and Icelandic switch to a mono-morphemic analysis of tense and agreement: an elsewhere form cannot be blocked by impoverishing the elsewhere morpheme that it spells out, because that morpheme does not encode any features that can be targeted by impoverishment. We are therefore bound to the assumption that the 3SG form is the elsewhere form. Abandoning it would allow us to capture the similar forms occurring in 1SG and 3PL contexts as syncretic, but we would lose the explanation for why Icelandic (and English) do not allow argumental pro drop.

The consequence for the Romanian fourth conjugation is that we have to postulate two distinct null forms and the question arises if this is not problematic for us. The answer is no, for three reasons. First of all, since the two contexts displaying the null forms do not form a natural class (and the 3SG context is not one of the contexts in which the null form appears, excluding the analysis of it being an elsewhere form), this counts as a case of unavoidable homonymy. Moreover, both null forms are paradigmatically licensed, the first one through the contrast with the two

other singular forms (2SG -*i* and 3SG -*e*), the second one through a contrast with the other plural forms (1PL -*m* and 2PL –*ţi).* Finally, neither null form is postulated as a way to capture a complementary distribution effect between two other forms. In the opposition of 1SG -∅ and 2SG -*i*, for instance, the 1SG -∅ form is the x in (14) and the 2SG -*i* form is the y. As stated before, (14) does not rule out that either x or y is a null form itself. Therefore, (14) has nothing to say about the null forms postulated in (22). They are therefore as innocent as postulating a null form expressing present tense.

4 The problem of the 3SG forms

        Let us now zoom in on a central problem of the Romanian paradigms, the observation that the form appearing in the 3SG of the present tense context does not return in the 3SG imperfect context. If imperfect tense is marked by -*a* and in the first conjugation the elsewhere agreement form -*ă* is inserted in 3SG imperfect contexts as well, we expect the outcome to be *cântaă* instead of *cânta*. For this reason, the 3SG person in the imperfect is sometimes taken as marked by a null form (cf. Zafiu 2012, 32). This form then functions as a null allomorph, as it replaces the -*ă* form we see in the 3SG present tense. Given (14), the prediction is then that a learner of Romanian would switch to a mono-morphemic analysis, like the English and Icelandic learner, with the unwanted consequence that we no longer expect pro drop to be possible. The question is therefore whether it is possible to rescue the bi-morphemic analysis.

        There are two logical possibilities we can pursue. One is to say that the forms appearing in the 3SG present tense contexts are in fact generated in the imperfect as well but that they get phonologically deleted. Let's call this *the deletion account*. In that way, we do not have to postulate a null allomorph and Romanian retains its bi-morphemic status. A second option is to say that the -*ă* and -*e* forms we see in the 3SG present tense contexts are not agreement markers but present tense indicative markers (Iliescu and Mourin 1991). If that is the case, we can take the 3rd person agreement marker to be null, like in Italian and Spanish, and this null form can then be taken to transparently return in the imperfect tense. In that case, there is no reason anymore to consider a mono-morphemic analysis. Let's call this *the null 3SG account*. We will first explore this latter option and then the first one. Both analyses, we argue, can be defended by reference to properties that play a role in Romanian at large. Although it is not entirely

obvious to us what the most viable analysis is, we speculate that the deletion account seems to pose fewer obstacles for language acquisition.

4.1 The null 3SG account

Let us look at the null 3SG account in detail. Following Illiescu and Mourin (1991), we could take 3SG -*ă* and -*e* in the present tense to be forms that denote the present indicative rather than person and number agreement. The core argument for this analysis is that in the subjunctive paradigm the 3SG is always marked by an opposite vowel. In the first conjugation, 3SG present tense is -*ă,* and the 3SG subjunctive -*e*, whereas in the other conjugations the 3SG present tense form is -*e* and we find -*ă* in the 3SG subjunctive. These contrasts would then justify the feature [indicative] for the 3SG markers in the present tense (or [Mood:] in a system that takes the indicative to be unmarked). It stands to reason to assume that these contrasts are part of the primary input data that children use to acquire Romanian because, in contrast to the other Romance languages, the subjunctive is used as an alternative to infinitival complementation, which makes it very pervasive in the input. The consequence of this analysis is that the 3SG agreement marker in the present tense is null, just like in Italian and Spanish:

(23)   -∅   <>   [Agr:]

We can subsequently assume that this null form is also generated in the past tense, as there is no evidence against it. The result of this is that there is no reason to switch to a mono-morphemic analysis because there is no agreement alternation to begin with, let alone one that requires postulation of a null form or homonym, as was the case in English and Icelandic.

In order to make this analysis work, we need to address a couple of issues. First of all, if -*ă* and -*e* are not agreement markers but markers of the present indicative, the question arises why we only see them occur in 3rd person contexts in the first conjugation, and only in 3SG context in the fourth conjugation. This question is especially pressing for the first person, where no overt marker follows the verbal root, as can be observed in (16). Second, if agreement is marked by a null form in the 3SG present tense context, it shares this property with the 1SG present tense context (cf. the analyses in (18) and (22). This is awkward in light of the fact that 1SG and 3SG contexts can be straightforwardly analyzed as a natural class: they are the singular contexts not marked by [addressee]. An analysis that

nevertheless postulates two distinct null forms for these contexts therefore postulates an avoidable homonym pair, and we have argued earlier that children avoid postulating such homonyms. On the other hand, if we treat the null form in the 3SG context as the same one that occurs in the 1SG context (in other words, we treat the null form as an elsewhere form that is also generated in the 1SG context by lack of an alternative), we beg the question why 1SG contexts do not show the hypothesized present indicative markers -*ă* or -*e*.

These problems can be addressed as follows. The reason why we only see the present indicative markers in a 3SG context is because in all other contexts it would be preceded or followed by a root-external vowel. We can therefore postulate the following deletion rule:[13]

(24)   -V$_{INDICATIVE}$ $\rightarrow$ $\emptyset$ / ..]$_{ROOT}$__V__

The triggering vowel can either be a following agreement vowel, as in 2SG -*i*, or a preceding theme vowel, like -*a* in the first conjugation. In the first conjugation, this accounts for the absence of the indicative marker in 2SG contexts (because of the adjacent agreement marker) and 1PL/2PL contexts (because of the adjacent theme vowel) and its presence in 3SG/3PL contexts, where agreement is hypothesized to be null and the theme vowel is absent. The remaining question is then why the indicative marker does not show up in 1SG contexts, where no obvious marker appears following the root. Focusing on the first conjugation, an answer is readily available, namely the presence of an underlying vowel. It can be argued that the 1SG agreement marker is -*u* but this marker only surfaces in specific phonological contexts, namely after a root-final consonant cluster consisting of an obstruent plus a liquid or after a root-final vowel, where it is realized as a /w/-glide (cf. Chitoran 2002). This is not something that has to be stipulated for 1SG verbs. The exact same contexts account for the realization of -*u* in a subset of masculine/neuter nouns and adjectives. So we have the following minimal pairs for verbs, nouns and adjectives, respectively.

(25)
a.   cânt(-u)   afl-u   bea-u
       sing.PRES.1SG   learn.PRES.1SG   drink.PRES.1SG
b.   lup(-u)   teatr-u   pusti-u
       wolf.M.SG   theatre.N.SG   desert.N.SG
c.   frumos(-u)   albastr-u   auri-u
       beautiful.M/N.SG   blue.M/N.SG   golden.M/N.SG