Isogloss
Open Journal of Romance Linguistics

Received: 13-10-2024
Accepted: 06-03-2025
Published: 14-07-2025

V-movement, subject clitics, and inversion

Edward J. Rubin
University of Utah
e.rubin@utah.edu

How to cite: Rubin, Edward. 2025. V-movement, subject clitics, and inversion. RLLT 25, eds. Adam McBride, James Law & Willis Fails. Special Issue of *Isogloss. Open Journal of Romance Linguistics* 11(5)/2, 1-20.
DOI: https://doi.org/10.5565/rev/isogloss.516

Abstract

Building on new evidence from Bolognese, this paper proposes an account for the appearance and distribution of φ-related subject clitics both within and among Gallo-Italic grammars that relies only on independently motivated mechanisms. It notes problems for previous accounts that include additional subject clitic specific functional heads and varying height of head-movement. Instead, it proposes that subject clitics depend on standard Agree, and that they are a consequence of an extra *u*φ ("agreement doubling") beyond the one inherited from C. Inversion in interrogatives is purely syntactic, and follows from strictly cyclic head-movement of V up to C. These two separate *u*φs underlie the distinct patterns of syncretisms and gaps in paradigms for verbal agreement suffixes and subject clitics in particular grammars, and for generalizations among them. Furthermore, gaps in a paradigm of subject clitics, like gaps in an agreement paradigm, do not imply a lack of Agree, but only a lack of overt exponents. Independently motivated syntactic mechanisms thus feed independently motivated morphological mechanisms. The syntax determines word/morpheme order, while morphophonology determines their realization.

Keywords: clitics, agreement, tense, interfaces.

1. Introduction

The subject clitics (SCLs) of the Gallo-Italic grammars have been intensively studied for a variety of goals, including what they can illuminate about the structure of clauses, the nature of agreement, the (Case-)licensing of subjects and many other related

phenomena. Microvariation among the many closely related grammars promises to clarify the finest details of the middle area of clausal structure where non-thematic but argument-related grammatical properties are encoded. This paper introduces data from (Urban) Bolognese that has not previously been considered in this generative context, and it proposes a novel analysis that refines earlier ones while avoiding various problems that they contained.

These SCLs derive from Latin strong pronouns with nominative case (see Vanelli 1987:184-5, among others), and are typically exemplified by verbal paradigms like (1), usually involving an intransitive verb such as *sleep*, *eat*, *drink*, *come*. Here, an optional tonic pronoun is included to demonstrate that overt subjects can co-occur with SCLs, which is typical of many of the relevant grammars. See also (8) below for an example with both an a SCL and a full r-expression in subject position.

(1) Bolognese, present tense, root: /do:rm/ 'sleep', Vitali (2007:260)

a.	(ˈme)	a=	ˈdo:rum	d.	(ˈno)	a=	ˈdurˈmæŋ
	I	SCL.1SG=	sleep		we	SCL.1PL=	sleep.1PL
b.	(ˈte)	t=	ˈdo:rum	e.	(ˈvo)	a=	durˈmi:
	you.SG	SCL.2SG=	sleep		you.PL	SCL.2PL=	sleep.2PL
c.	(ˈlo/ˈli)	al/la=	ˈdo:rum	f.	(ˈlɑwr)	i/æl=	ˈdo:rmeŋ
	he/she	SCL.3SG.M/F=	sleep		they	SCL.3PL.M/F=	sleep.3PL

Declarative data like (1) is often contrasted with interrogative data like in (2), and in many Gallo-Italic grammars (but not all, as shown in surveys of multiple grammars like Renzi & Vanelli (1983) and Manzini & Savoia (2005)), *inversion* is observed, in the sense that the SCL occurs before the inflected form in the declarative form, as a proclitic, but after it, as an enclitic, in the interrogative. (Inversion may occur in other constructions, but we put that aside, here.) As we will discuss below, many researchers hold that inversion relates to the notion of head-movement, T-to-C movement in particular.

(2) Bolognese, present tense with inversion, Vitali (2007:260)[1]

a.	ˈdo:rm=ja ?	d.	durˈmaɲa ? (< /durmæn=ja/)
	sleep=SCL.1SG		sleep.1PL=SCL.1PL
b.	ˈdo:rm=et ?	e.	durˈmi:=v ?
	sleep=SCL.2SG		sleep.2PL=SCL.2PL
c.	ˈdo:rm=el/la ?	f.	ˈdo:rmn=i ?
	sleep=SCL.3SG.M/F		sleep.3PL=SCL.3PL

[1] The underlying form in (2d) is based on Canepari & Vitali (1995), who show that surface forms like this involve the combination of /n/ and /j/ into [ɲ]. They furthermore show that [æ] generally only survives in modern Bolognese before [ŋ], as in (1d) (but also in the preverbal SCL.3PLF and its homophonous definite determiner). Similarly, the [n] in (2f) is the allomorph for 3PL verbal agreement. Canepari & Vitali (1995) and Vitali (2007) show that Bolognese /n/ becomes [ŋ] in codas, as in (1f) and throughout the grammar. Rubin & Kaplan (2024) demonstrate that the [e] in (1f), those preceding the enclitics in (2b,c), as well as the [u]s that appear in (1a-c) but not in (2a-c), are predictably epenthetic.

A deeper understanding of alternations like these helps in avoiding claims that there are unpredictable differences between proclitic and enclitic forms. Many paradigms available in the literature might benefit from careful study of the sound patterns of the grammar under investigation.

We must note an important distinction in the literature concerning paradigms like this. In the 1SG, 1PL, and 2PL cells of data like (1), the preverbal forms are often homophonous, [a] in Bolognese (though the specific form differs among grammars). Some works treat homophonous forms like these three as related to the φ-features of the subject, as glossed here (e.g. Rohlfs 1968, Renzi & Vanelli 1983, Manzini & Savoia (2005), Roberts 2010, Calabrese 2011, etc). Other works do not, treating them distinctly from the φ-related SCLs in the 2SG, 3SG, and 3PL cells (e.g. Benincà 1983, Poletto 2000, Cardinaletti & Repetti 2008, etc). Works like these latter ones often provide convincing evidence that, in the grammars that they investigate, these elements show properties that differ from those of φ-related SCLs. It is sometimes implied that φ-related SCLs are impossible in these cells, but this is not necessarily so. It is possible that some grammars have these non-φ-related elements and no (realized) φ-related SCLs in some cases, and that other grammars have φ-related SCLs in all cells, which is implied by their noted evolution from tonic pronouns.

In fact, Benincà's (1983) original discussion of separating φ-related SCLs from other elements suggests that the latter are a further step in diachronic development of the original Latin pronouns. Benincà (1983:25) discusses the Ruzante variety of Paduan and claims that it does not show the evidence for distinguishing the homophonous elements in 1SG, 1PL, and 2PL cells from the φ-related SCLs in the other cells, stating that "in the subsequent century" the relevant distinctions arose. Before the transition, these elements patterned identically to the SCLs that are generally analyzed as involving agreement features. Although the observable form of the homophonous elements does not change, after the transition, it is argued that these elements no longer associate with agreement features, but instead serve some other function. Benincà's claims support the notion that grammars may indeed have a full set of φ-related SCLs, since Ruzante is such a grammar. As discussed below, Bolognese is another one. Each grammar must be evaluated in light of whether these elements pattern with other φ-related SCLs or otherwise.

A related point arises from Calabrese's (2011) morphological treatment of the homophony among these SCLs (or sub-sets of them) in a large set of grammars as due to syncretism (or obliteration, discussed below). Some of these grammars may in fact have elements serving other functions in them, rather than homophonous φ-related SCLs, though the needed analysis has not been done for most of these grammars. Nevertheless, it is entirely possible to hold Calabrese's analysis of those grammars as applying to an earlier historical moment in their diachronic development, before the transition of these elements from one type to the other. This seems likely, given the discussion in Benincà (1983), and we can take this sort of pattern of development to have applied to all grammars in which it can be argued that non-φ-related "SCLs" occur, such as Paduan.

Importantly, though, this idea may undermine a wide-spread notion in the literature that observed variations in patterns of overt SCLs is due to purely syntactic factors. Instead, it may derive, as Calabrese (2011) argues, from morphological factors. It must be admitted that grammars permit allomorphy (e.g. [s] vs. [z] in English *sleeps* vs. *sings*), and that they allow null allomorphs (e.g. the 1SG agreement in English *I sleep*). We should note in this regard that many grammars with SCLs lack either a φ-related SCL or non-φ-related element in some cells of the (predominately PRES tense) paradigms that have been studied. This might be understood as gaps in the paradigms for the relevant SCL morphemes, similar to the ones in English agreement, rather than the complete absence of that morpheme. Renzi & Vanelli (1983) proposed

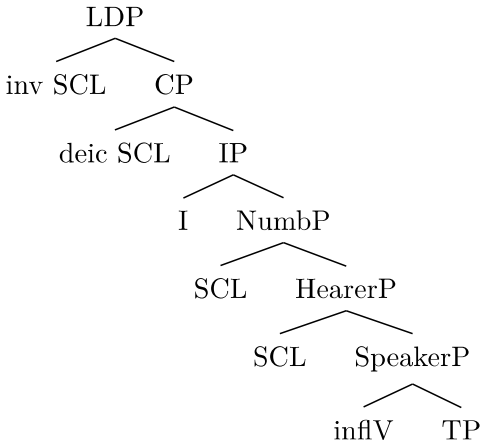
cross-grammar generalizations regarding the order of appearance of SCLs in particular grammars based on these gaps in paradigms. Calabrese's proposal indicates that these differences in paradigmatic gaps may be better explained by differential morphological obliteration (constrained by markedness of agreement features) rather than by differences in the syntax (head-movement, etc) of the relevant grammars. As discussed below, leaving the explanation of these generalizations to morphological mechanisms as well as details of the synchronic morphemes in each grammar allows us to propose a simple and parsimonious analysis of the syntax underlying all observed data.

2. Pros and cons of previous analyses

In this section, we review some previous analyses for SCLs in the Gallo-Italic grammars. We observe aspects of each of them that provide useful insight as well as some difficulties with the sort of deeper analysis of specific grammars provided here for Bolognese, and of other data presented in the literature.

Poletto (2000), building on Benincà's (1983) insights discussed above, proposes an account of SCLs within the cartographic approach that differentiates among four types and requires four additional functional heads, as seen in (3):

(3) Poletto (2000:36)



In this approach, the proclitic φ-related SCLs include only the 2SG, 3SG, and 3PL ones, and the homophonous pre-verbal elements in the 1SG, 1PL, and 2PL cells of data like (1) are argued to "not encode any subject feature at all" (Poletto 2000:12), to be discourse-related elements, and to occur in the head of LDP (named for its association with left-dislocated elements in certain data types. See the discussion in Poletto 2000:Chapter 2 for details). For the grammars specifically investigated, the evidence that these elements are discourse-related is convincing.

In contrast, however, the types of data that support the analysis in (3) for the grammars investigated in Poletto (2000) do not occur throughout the grammars that have φ-related SCLs and inversion. In Bolognese, none of these tests for being high discourse-related elements, e.g. the co-occurrence of distinct types in (4), produce similar results to those in the grammars studied there. Indeed, they do not distinguish among Bolognese SCLs at all, and they all appear to show properties of being φ-related, just as Benincà (1983) argued for the Ruzante variant of Paduan in comparison to the

more recent variant of Paduan in (4c). Despite having additional specific SCL-related functional heads, it is also not clear that (3) provides positions (or a position) for the homophonous proclitic φ-related 1SG, 1PL, and 2PL SCLs of Ruzante or Bolognese.

(4) Montagnola, Poletto (2000:20)
a. A ta vegnat.
SCL SCL come
'You come.'
Bolognese, field work[2]
b. (*A) t veŋ.
SCL.2SG= come
Padua, Poletto (2000:20)
c. A i vien.
SCL SCL come.
'They are coming.'
Bolognese, field work
d. (*A) i veŋneŋ.
SCL.3PL= come

In this regard, Poletto (2000:30) notes that "there are no SCLs that morphologically distinguish first person SCLs," and on the next page argues that (3) explains this, since the inflected verb occupies the head of SpeakerP, which is where a first person SCL might be realized. It is not clear, however, that a lack of morphological distinction in certain feature sets (gaps in a paradigm) implies the lack of an underlying morpheme that requires a syntactic explanation, as discussed above.

Given differences in forms between proclitics and enclitics, Poletto (2000) provides a distinct analysis for the latter, not treating them as realizations of the same heads. She makes the standard assumption that the inflected verb raises higher in structures with enclitics than in structures with proclitics. This is essentially T-to-C movement, where, according to this analysis, T is equated to Speaker in (3). Since proclitics are realizations of the heads between Speaker and the higher position to which V raises, they cannot occur when V raises through them. Roberts (2010:117) says this analysis "treats the enclitics as a realization of interrogative conjugation." Cardinaletti & Repetti (2008:525) refer to this sort of distinction between proclitics and enclitics as "the two-paradigm hypothesis." Like these researchers, the current analysis rejects this notion and derives both proclitics and enclitics from a common underlying morpheme, as discussed below.

However, we must note that Poletto's analysis treats proclitics as distinct from the inflected verb, that is, the two items do not form a constituent. This is different from the other analyses to be discussed that overtly treat them as a single constituent derived by incorporation or head-adjunction of the clitic to the inflected verb. Our discussion and proposal below shall provide evidence that a non-constituent analysis of proclitic SCLs and inflected verbs similar to Poletto's is preferrable.

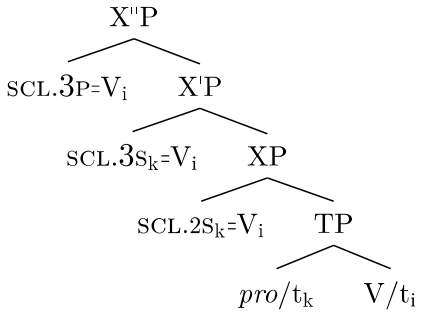
A second important analysis of SCLs is presented in Cardinaletti and Repetti (2008), which, as noted, takes both proclitics and enclitics to derive from the same morpheme, and handles variations in forms (but not gaps in paradigms) in the morphophonology. In this analysis, cliticization involves the syntactic process of head-adjunction, moving the SCL to a head from the specifier of that head's complement, following Cardinaletti & Starke (1999).

For proclitics, their proposed structure is shown in (5). Like Poletto (2000), it assumes additional functional heads associated with the 2SG, 3SG, and 3PL φ-related SCLs. As an example, a SCL.3S would incorporate into the head of their X'P in

[2] Data marked as "field work" come from questions asked of consultants in and around Bologna, or from observations/recordings of their conversations/presentations/recitals, or from questionnaires to which they responded, or from their public videos on the internet.

grammars in which the verb can raise to X', like Donceto (which they investigate), or Bolognese, or others. In grammars that lack these φ-related SCLs, they hold that the verb cannot raise as far as the corresponding head, and that the subject is realized as the weak pronoun *pro*.

(5) Cardinaletti and Repetti (2008:550)



Also like Poletto (2000), they provide evidence that the preverbal material in 1SG, 1PL, and 2PL data ([ə] in Donceto) is not a φ-related SCL, but rather the default realization of a functional head. Their proposal implies that 1SG, 1PL, and 2PL φ-related proclitic SCLs cannot exist, since there aren't heads in their proposal associated with those features to which a verb could raise. In such data, the verb always remains below, in T, with *pro* in its specifier, thus precluding head-adjunction and an overt SCL.

For enclitics, T-to-C movement occurs, with C above X"P in (5), and the configuration for head-adjunction is made more generally available. This accounts for the greater number of enclitics than proclitics in Donceto and other grammars, which is one of Renzi & Vanelli's (1983) generalizations.

There are some difficulties with this analysis. As they note, "further crosslinguistic investigation is of course needed to establish exactly what triggers verb movement to the X head(s) in the different NIDs. (pg. 550)" Moreover, head-movement is clearly not the only factor involved in dealing with grammars that have only a partial set of SCLs: They say that in such cases, "no pronoun is available (pg. 547)." This is indistinguishable from cases involving *pro* where a pronoun/SCL is available for the enclitic. (This point is, however, reminiscent of the notion of morphological obliteration, discussed above, and we will expand on it in our analysis below.) Finally, although enclitics are derived syntactically via the same head-adjunction as proclitics, their ordering relative to the head to which they adjoin is manifestly the opposite. For proclitics, it is simple left-adjunction. For enclitics, they adopt the morphological mechanism of "completeness" from Rizzi (2000), and so, again, head-movement and syntactic adjunction is not sufficient to characterize the facts in this account.

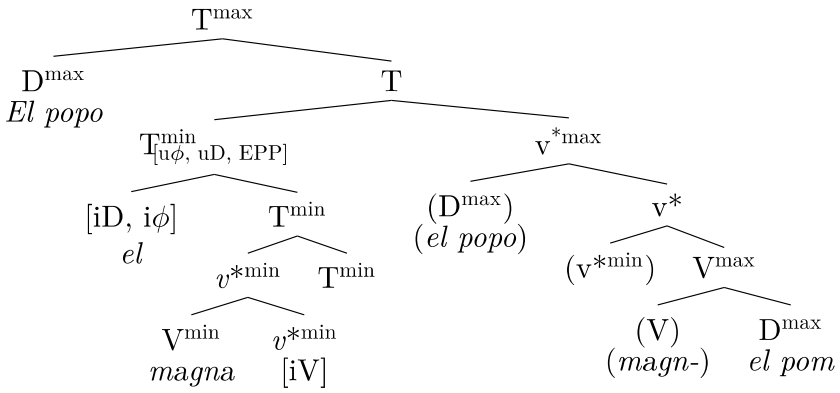
Another important previous account to consider is the one found in Roberts (2010), exemplified for the data containing a proclitic in (6) by the tree in (7). (Bolognese (8), which is essentially identical to (6), might receive the same analysis.)

(6) Montesover Trentino, Roberts (2010:109), from Manzini and Savoia (2005)

El	popo	*(el)	magna	el	pom
the	child	he	eats	the	apple

'The child eats the apple.'

(7) Montesover Trentino, Roberts (2010:109)



(8) Bolognese, field work
al faŋˈʤeŋ *(al) ˈmaːɲa la ˈmajla
the child SCL.3MSG= eat.3SG the apple
'The child eats the apple.'

Importantly, Roberts (2010:106) invokes the notion of "agreement doubling" from the dialectological literature, and notes that the SCLs together with the normal agreement inflection suffix on the verb often mark agreement distinctions that they each fail to mark individually. He notes that syncretisms and gaps reduce distinctions in each expression of agreement with a subject like the ones in (6) or (8), or with all the subjects in a paradigm like (1). For these null-subject languages, he holds that the combinations of agreement and a SCL "can distinguish at least five persons. (pg. 108)" This is true in (1), where all six forms are distinguished in interpretation even when the tonic pronoun is absent, despite the homophony of the singular forms (which in fact all have null agreement morphemes). (We must also note that not all Bolognese verb paradigms fall neatly into only one of Roberts' (2010:106) four types of grammars with SCLs, as we shall see in the data in (10-17). There are always four formally distinct proclitic SCLs, ignoring gender distinctions, including the one syncretism [a], and five formally distinct enclitic SCLs. However, the number of distinctions in agreement suffixes varies from four to six, with some paradigms showing more than one distinct syncretism. Nevertheless, all paradigms allow complete distinction of all six forms.)

Roberts (2010) formalizes this agreement doubling by holding that both the agreement suffix and the SCLs are manifestations of φ. The agreement suffix is *u*φ inherited by T from C, and the SCLs are *i*φ incorporated from v*^{max} into T from the subject A-chain in a manner identical to the head adjunction proposed in Cardinaletti & Repetti (2008). He does not discuss whether or not the A-chain subject also retains *i*φ, and whether or not this issue creates problems. We put this issue aside and, in the proposal below, we will take up Roberts' notion that SCLs are the manifestation of an additional φ beyond the agreement suffix and any expression of φ in the A-chain subject.

For enclitics, Roberts' (2010:118-9) analysis becomes slightly different, and connects to his analysis of residual V2. Instead of the incorporation of *i*φ into T due to its *u*φ inherited from C, *u*φ remains on C, and enclitic SCLs "are thus weak pronouns, first merged in SpecvP, raised to SpecTP, and then cliticized to the low C-position."

Finally, let us consider the data in (9), which presents an important empirical challenge to certain types of analyses of SCLs.

(9) Triestino, Benincà (1997), Ledgeway & Lombardi (2005)
a. el me sempre dizi
'he always tells me.'
Triestino, Paoli (2008)
b. Ma te lo pena conossi!
but you=him=barely you-know
'But you barely know him!'

It was first noted by Benincà (1997) that Triestino has SCLs as well as the phenomenon that Ledgeway & Lombardi (2005) called *interpolation*, where certain adverbs can intervene between a clitic and the inflected verb. Ledgeway & Lombardi (2005) provided an analysis for interpolation in the southern Italian grammar of Cosentino, which is not Gallo-Italic and lacks SCLs. Paoli's (2008) analysis of Triestino picked up the core of their proposal, that interpolation occurs when the inflected verb does not raise as high as the position of the (object) clitic.

Two important conclusions can be drawn, (i) that the inflected verb in Romance may not always raise overtly as far as T, even when, as in Triestino, a SCL is present, and (ii) that clitics like those in (9) do not necessarily form a constituent with the inflected verb. As noted earlier, the analysis in Poletto (2000) holds that the inflected verb does not raise as far as proclitic SCLs and that the two do not form a constituent, and in this regard can be extended to Triestino. On the other hand, the analyses in Cardinaletti & Repetti (2008) and Roberts (2010) crucially depend on the two becoming a constituent through high head-movement of the inflected verb and head-adjuction/incorporation of the clitic to it. As a result, such analyses of SCLs cannot be straightforwardly extended to grammars like Triestino with both SCLs and interpolation.

3. Grammar internal data variation

Most of the research on SCLs focuses on paradigms of forms in the present tense, like those that we have seen in the previous data. If, however, we investigate verbal forms that express other tense forms, like the Imperfect (IMPF) forms in (10-11), we find that the patterns involving SCLs are, for the most part, the same as those in present tense paradigms, but there is (at least) one important and revealing difference. The preverbal SCLs in (10) show the same forms as we saw in (1). However, in the inverted forms in (11), we see that the SCLs are mostly as expected, except in the 2PL (11e), where there is no overt SCLs at all (as indicated by [=Ø]).[3]

[3] Summarizing the facts observed so far, Bolognese SCLs show the following forms:

(i)
a. [a]/[ja] SCL.1SG
b. [t] SCL.2SG
c. [al]/[l] SCL.3SGM; [la] SCL.3SGF
d. [a]/[ja] SCL.1PL
e. [a]/[v]/[Ø] SCL.2PL
f. [i] SCL.3PLM; [æl] SCL.3PLF

(10) Bolognese, imperfect forms, root: /do:rm/ ‘sleep’, Vitali (2007:260)

a.	a=	durˈmɛːva	d.	a=	durˈmɛːveŋ
	SCL.1SG=	sleep.IMPF		SCL.1PL=	sleep.IMPF.1PL
b.	t=	durˈmɛːv	e.	a=	durˈmɛːvi
	SCL.2SG=	sleep.IMPF		SCL.2PL=	sleep.IMPF.2PL
c.	al/la=	durˈmɛːva	f.	i/æl=	durˈmɛːveŋ
	SCL.3SG.M/F=	sleep.IMPF		SCL.3PL.M/F=	sleep.IMPF.3PL

(11) Bolognese, imperfect forms with inversion, Vitali (2007:260)

a.	durˈmɛːv=ja ?	d.	durˈmɛːvŋa ? (< durˈmɛːvn=ja)
	sleep.IMPF=SCL.1SG		sleep.IMPF.1PL=SCL.1PL
b.	durˈmɛːv=et ?	e.	durˈmɛːvi=Ø ?
	sleep.IMPF=SCL.2SG		sleep.IMPF.2PL=SCL.2PL
c.	durˈmɛːv=el/la ?	f.	durˈmɛːvn=i ?
	sleep.IMPF=SCL.3SG.M/F		sleep.IMPF.3PL=SCL.3PL

Analyses of the appearance and distribution of enclitic SCLs that depend on head-movement of the verbal form to a position higher than the position of the subject have difficulty with this last type of observed data, where there is an alternation between an overt and a null enclitic. That is, in PRES data like (2), such accounts hold that overt forms of the enclitic result from head-movement of the verbal form to a position higher than the subject, in the left periphery, and this movement then results in the overt [v] form in (2e), just as it does for the other five forms. The morphemes are available, and they arise when this configuration occurs. In IMPF data like (11), however, the same head-movement should result in the same outcome, and the outcomes are in fact the same in all but this one case. There is no good reason to suppose that head-movement applies equally in all the data in (2) and in five of six cases in (11), but not in (11e) where the subject is 2PL. This combination of a tense feature with particular φ-features of the subject (or with the *u*φ that results in the separate overt verbal agreement morpheme) should no more interfere with head-movement to C than any other combination might. We must conclude that head-movement to C alone cannot explain the presence or absence of this enclitic SCL.2PL. The implication is that this head-movement alone cannot determine gaps in SCL paradigms within a grammar, nor that it alone can explain cross-grammar patterns of gaps. We return to this issue below in our proposal.

It would certainly be useful to investigate a wider range of grammars in this regard, looking beyond just the PRES data that is normally considered. The Bolognese pattern regarding the SCL.2PL (which is difficult for previous proposals to explain) is in fact quite general and might also be so in many grammars. The Bolognese future (FUT) form data in (12-13) patterns identically with the PRES data in (1-2), showing a proclitic SCL.2PL [a] and an enclitic SCL.2PL [v]. On the other hand, the past (PST) data in (14-15) and the conditional (COND) data in (16-17) patterns identically with the IMPF in (10-11), showing a proclitic SCL.2PL [a] but no overt enclitic form at all. (These

Recall, as discussed above, the [e] observed with the enclitics of SCL.2SG and SCL.3SGM is predictably epenthetic. Not shown in this paper, the SCL.3PLM is realized as [j] prevocalically.

groupings are perhaps unsurprising, given the typical Romance origin of FUT forms in a Latin PRES, and of IMPF and COND and PST forms in Latin PST forms.)

(12) Bolognese, future forms, root: /do:rm/ ‘sleep’, Vitali (2007:260)

a.	a=	durmiˈrɔ:	d.	a=	durmiˈræŋ
	SCL.1SG=	sleep.FUT.1SG		SCL.1PL=	sleep.FUT.1PL
b.	t=	durmiˈrɛ:	e.	a=	durmiˈri:
	SCL.2SG=	sleep.FUT.2SG		SCL.2PL=	sleep.FUT.2PL
c.	al/la=	durmiˈra	f.	i/æl=	durmiˈraŋ
	SCL.3SG.M/F=	sleep.FUT.3SG		SCL.3PL.M/F=	sleep.FUT.3PL

(13) Bolognese, future forms with inversion, Vitali (2007:260)

a.	durmiˈrɔ:=ja ?	d.	durmiˈraŋa ?
	sleep.1SG=SCL.1SG		sleep.1PL=SCL.1PL
b.	durmiˈrɛ:=t ?	e.	durmiˈri:=v ?
	sleep.2SG=SCL.2SG		sleep.2PL=SCL.2PL
c.	durmiˈrɛ:=l/la ?	f.	durmiˈrɛ:n=i ?
	sleep.3SG=SCL.3SG.M/F		sleep.3PL=SCL.3PL

(14) Bolognese, past forms, root: /do:rm/ ‘sleep’, Vitali (2007:260)

a.	a=	durˈme	d.	a=	durˈmen
	SCL.1SG=	sleep.PST		SCL.1PL=	sleep.PST.1PL
b.	t=	durˈmes	e.	a=	durˈmesi
	SCL.2SG=	sleep.PST		SCL.2PL=	sleep.PST.2PL
c.	al/la=	durˈme	f.	i/æl=	durˈmen
	SCL.3SG.M/F=	sleep.PST		SCL.3PL.M/F=	sleep.PST.3PL

(15) Bolognese, past forms with inversion, Vitali (2007:260)

a.	durˈme=ja ?	d.	durˈmeŋa ?
	sleep.PST=SCL.1SG		sleep.PST.1PL=SCL.1PL
b.	durˈmes=et ?	e.	durˈmesi=Ø ?
	sleep.PST=SCL.2SG		sleep.PST.2PL=SCL.2PL
c.	durˈme=l/la ?	f.	durˈmen=i ?
	sleep.PST=SCL.3SG.M/F		sleep.PST.3PL=SCL.3PL

(16) Bolognese, conditional forms, root: /do:rm/ ‘sleep’, Vitali (2007:260)

a.	a=	durmiˈre:v	d.	a=	durmiˈreŋ
	SCL.1SG=	sleep.COND		SCL.1PL=	sleep.COND.1PL
b.	t=	durmiˈres	e.	a=	durmiˈresi
	SCL.2SG=	sleep.COND		SCL.2PL=	sleep.COND.2PL
c.	al/la=	durmiˈre:v	f.	i/æl=	durmiˈreŋ
	SCL.3SG.M/F=	sleep.COND		SCL.3PL.M/F=	sleep.COND.3PL

(17) Bolognese, conditional forms with inversion, Vitali (2007:260)

a. durmiˈre=ja ?
sleep.COND=SCL.1SG

b. durmiˈres=et ?
sleep.COND=SCL.2SG

c. durmiˈre=l/la ?
sleep.COND=SCL.3SG.M/F

d. durmiˈreŋa ?
sleep.COND.1PL=SCL.1PL

e. durmiˈresi=Ø ?
sleep.COND.2PL=SCL.2PL

f. durmiˈren=i ?
sleep.COND.3PL=SCL.3PL

We must conclude that, in Bolognese at least, there are at least three allomorphs of the SCL.2PL, [a=], [=v], and [=Ø].

Data like (10-17) are not generally available or discussed in the generative literature concerning SCLs, nor even in the treatments of specific grammars. For example, such forms are not generally available in Rohlfs (1966-9), Poletto (2000), nor in Manzini & Savoia (2005), to which we return below. Cardinaletti & Repetti (2008) present only PRES data, and Zörner's (1989) book length treatment of two other nearby western Emilian grammars presents full paradigms for forms with proclitics, but no corresponding paradigms with inversion/enclitics. Some few isolated examples in Zörner (1989) show an enclitic SCL.2PL [v] with PRES tense, but no examples show other tense forms.

Manzini & Savoia (2005:360-379) provide data from 93 grammars with inversion (the Bolognese data presented here expands that set), from "the interrogative paradigm of the present indicative of *to sleep* or *to eat* … and possibly [from] the paradigm of the imperfect or conditional indicative." In their paradigms, there is data from tenses other than PRES in only six of the 93 grammars. Table 1 presents all their 2PL forms contrasting PRES forms with IMPF or COND ones, using forms of 'sleep' or 'eat'.

Table 1. Variations in SCL.2PL forms

	Proclitic		**Enclitic**	
Grammars	**PRES**	**IMPF [COND]**	**PRES**	**IMPF [COND]**
Pozzaglio/ Cicognòlo	durˈmi	durˈmivɛs	durˈmi:?	durˈmivɛs?
Castellinaldo	(e) ˈmaɲˈdʒi	(e) maɲˈdʒɔvi	maɲˈdʒa-v-i?	maɲˈdʒɔvi?
Ro Ferrarese	a durˈmi	a durˈmivi	durˈmi:f?	durˈme-vi?
		[a durmiˈresi]		[durmiˈres-i?]
Viano	durˈmi	(NA)	durˈmi-o?	əl vəˈdeo-vo?
Corte	maɲˈdʒei	maɲˈdʒeive	maɲˈdʒei-zo?	maɲˈdʒeiv-io?
Montereale	durˈmi	maɲdʒeˈai	durˈmi-va?	maɲdʒeˈai-iu?
		[maɲdʒeˈsai]		[maɲdʒeˈre-iu?]

Source: Manzini & Savoia (2005)

Despite Manzini & Savoia's (2005) impressive work, we must all admit that further documentary and analytical work is required in many grammars, as the differences in various tense forms may indicate a greater range of allomorphs of the SCLs than has previously been considered when positing analyses of their distribution. In the following section, a new analysis is proposed that captures the important empirical

issues discussed above, and that draws on the best aspects of previous works, while avoiding their problems.

4. A new proposal

At the core of this proposal is the previously discussed notion that φ-related SCLs are the manifestation of agreement doubling, in the sense that they reflect a second realization of φ-features that depend on the subject. In contrast to Roberts' (2010) implementation of this idea via incorporation of the interpretable φ-features of that subject, as seen in (7), we propose that grammars with SCLs introduce a second *u*φ into derivations that is distinct from the standard *u*φ associated with verbal agreement morphology (overt or null) and with the Case-licensing of subjects.

Since the realization of particular SCLs is as arbitrary as any other (inflectional) morpheme (though of course related to its particular diachrony), we follow Cardinaletti & Repetti (2008) in taking SCLs to be morphemes, and furthermore in taking both proclitic and enclitic allomorphs to derive from the same morpheme. The observed variation among Gallo-Italic grammars occurs as a normal result of (micro-varying) historical change. Following Calabrese (2011), we hold that syncretisms and gaps in the paradigms of SCLs are handled as in other paradigms, by morphological considerations rather than by differences in overt syntactic movement. Importantly, gaps in a paradigm, for example the lack of overt allomorphs of proclitic φ-related 1SG, 1PL, and 2PL SCLs in Donceto or (recent) Paduan, or any other gaps in the paradigms of other grammars with SCLs, do not imply the absence of the additional *u*φ that occurs in these grammars or of the Agree that values and deletes it, and that furthermore allows for the determination of the allomorph, overt or null, that realizes it after Spell-out. Indeed, it is this same additional *u*φ that, because of Agree, allows the realization of this morpheme in enclitic contexts.

Finally, we follow all previous work in holding that T-to-C movement underlies inversion, but we reject the notion of novel heads whose sole function is to host specific proclitic SCLs. Instead, we only adopt heads that are independently motivated.

Recall that there is strong evidence in support of the notion that proclitics and the inflected verb do not form a constituent. In a grammar like Triestino (9), the inflected verb may be separated from clitics by an adverb, and the inflected verb must be realized in a position below the position of the SCL. However, the height of raising of the inflected verb can vary across grammars, and we shall exemplify our proposal using structures similar to those in Poletto (2000), where the inflected verb has overtly raised to a position close to that of the proclitic SCLs, without precluding that it may be lower in the structures of other grammars, including Triestino.

For the specific structure in which SCLs occur, we adopt the one in Rizzi (2015:26) that includes the head Subj(ect) (which has the LF interpretation of "aboutness", noted as *i*ABOUT): "… Fin … Subj … Phi … T ……". The Subj head is motivated by extensive discussion that is not related to SCLs, and it is thus an independently motivated head. Furthermore, he says, "Phi may or may not be identified with T," and we leave it out of subsequent trees both for space and because it is not necessary for our proposal.

Rizzi (2015:26) moreover does suggest that "a system of subject clitics distinct from agreement morphology may overtly instantiate the Subj head," and cites the data

in (18), which is essentially identical to (6) and (8). He provides no further details of such a claim, but we incorporate his point into our elaborated proposal.

(18) Milanese, Poletto (2000); Manzini & Savoia (2005), etc.
El fio el mangia l pom
'The boy SUBJ eats the apple'

We can now present analyses in terms of our proposal. We shall provide trees for the data of particular interest, starting with proclitics in the PRES and IMPF tenses, and following those with trees for data including enclitics in those tenses. It should be noted that these trees can be generalized to data involving the other synthetic forms, the FUT (12-13), PST (14-15), and COND (16-17). We focus here on data with a 2PL thematic subject, since it is in this φ-combination that the most surface variation in forms is observed.[4] Despite this variation, the same essential syntax will apply throughout, with the only difference being in the independently motivated T-to-C movement that occurs with data including enclitic SCLs, but not with proclitic SCLs.

Let us start with the data in (19), which is similar to others discussed above and in the literature. Data like this is represented as in (20) according to our proposal.[5]

(19) Bolognese, field work

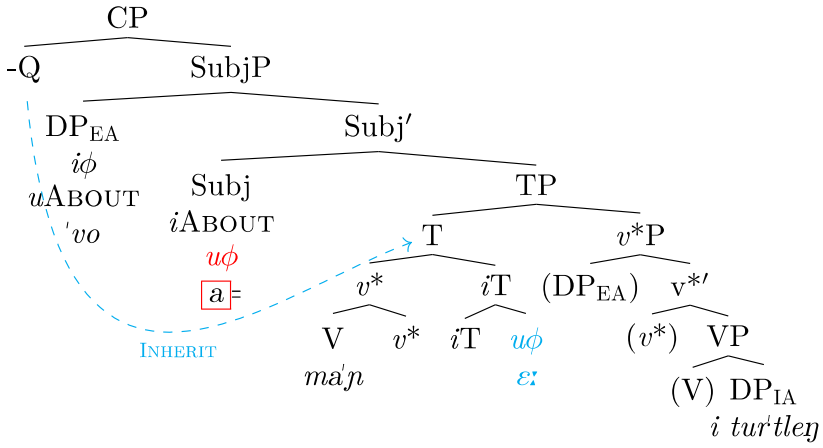
ˈvo	a	maˈɲɛː	i	turˈtleŋ
you	SCL.2PL	=eat.2PL	the	tortellini

'You are eating the tortellini.'

[4] For reasons of space, a full formal analysis of the paradigms realized in Subj and in T is deferred to future work, in development. Similarly full formal analyses of the equivalent paradigms in other grammars is also deferred, both for space and because much of the relevant data is not documented, as discussed. For all of these, however, (a combination of) typical morphological and phonological factors must be considered, and as we shall see most explicitly, under this analysis the category of the word and its φ-features are relevant.

[5] An anonymous reviewer correctly points out that questions arise both "regarding the strong subject, the SCL, and the inflected verb" as they relate to potentially intervening parenthetical expressions, discussed in Cardinaletti (2004), and regarding the multiple subject positions posited in that work. Rizzi's (2015) discussion of the structures adopted from that work are relevant here, and can underlie a characterization of the relevant data by means of the expanded treatment of the structures involving SCLs adopted here. A full treatment of the positions available to various overt (strong) subjects in grammars with SCLs, however, is beyond the scope of this paper. Here we focus instead on the SCLs themselves, which Agree with subjects in their theta-positions, as described below.

(20) Proposed Declarative, Present Tense



The standard *u*φ occurs on T (through inheritance, or otherwise), and probes and Agrees with the highest argument DP, here the thematic subject (DP$_{EA}$). It becomes valued in this case as 2PL, which, after Spell-out, realizes the verbal agreement suffix as [ɛː].[6] Following Rizzi (2015), the thematic subject raises in such data to SpecSubjP, which accords with the observed data. According to our proposal, since Bolognese has SCL morphemes, the head Subj is endowed with an additional distinct *u*φ, which also probes and Agrees with the same highest argument DP, again being valued as 2PL. In this case though, it is realized after Spell-out as a SCL and not as an agreement suffix, since it is in Subj, and not in T. Here, it appears as the [a] allomorph of the SCL.2PL. Finally, note that the SCL in Subj does not form a constituent with the inflected verb in T. As discussed above, this makes our analysis compatible with the interpolation data in Triestino (9), assuming an even lower position of the inflected verb in that data.

Note that other grammars may have different realizations both for a *u*φ in T valued as 2PL, and, more importantly for our focus here, for *u*φs in Subj valued as 2PL. For example, in grammars like Paduan and Donceto, in this context, the realization is null, since an old φ-related form in this context had been reanalyzed diachronically as a new element with distinct properties, just as in Benincà's (1983) comparison of Ruzante and Paduan. The same null allomorph analysis probably also applies in grammars including Viano, Corte, and Montereale, seen in Table 1. In the PRES data there, there is no pre-verbal element at all, so no discussion of a type is necessary.

Next, we apply our analysis to the data in (21), which involves IMPF rather than PRES tense. As we can see, the only difference between its tree in (22) and the one in (20) is that (22) below has an overt morpheme for IMPF tense rather than the typically null morpheme for PRES tense above. Otherwise, our proposal applies exactly the same syntactic analysis as above, involving the two *u*φs and the effects of Agree on them. The result here is the same allomorph [a] in a context which certainly does seem identical to the one above.

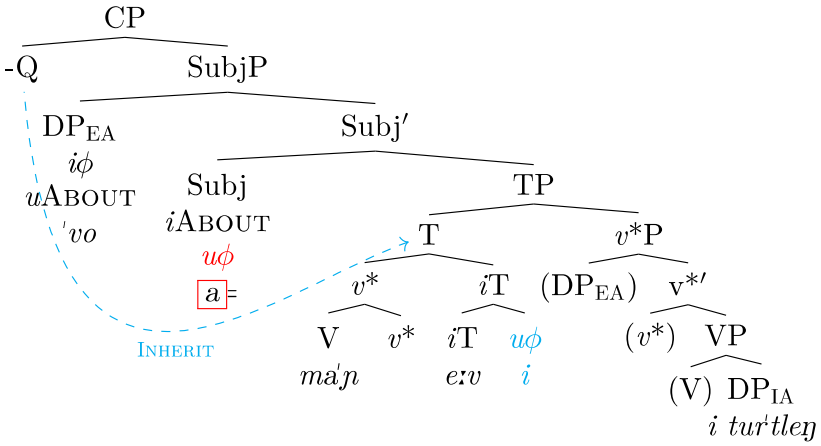
[6] Note that /maɲ/ 'eat' is first conjugation in Bolognese, while /doːrm/ 'sleep' in (1-2) and (10-17) is fourth conjugation, which would have the allomorph [iː] in this case. Many morphological differences of this normal type in Romance could possibly confound analyses of data involving SCLs, including the data in Table 1.

(21) Bolognese, field work

ˈvo	a	maˈɲeːvi	i	turˈtleŋ
you	SCL.2PL=	eat.IMPF.2PL	the	tortellini

'You were eating the tortellini.'

(22) Proposed Declarative, Imperfect Tense



Furthermore, the same null allomorphs occur for the IMPF data from Corte and Montereale seen in Table 1 as occur in their PRES data, further supporting our proposal's generalization of this structure and its idea that context determines the observed forms.

Next, we turn to data involving inversion and enclitic SCLs. We start with the PRES forms and follow it with the IMPF forms. Once again, we find that our proposal applies identical syntax in all cases, but with the addition here of independently motivated T-to-C movement that is absent in the proclitic examples above.

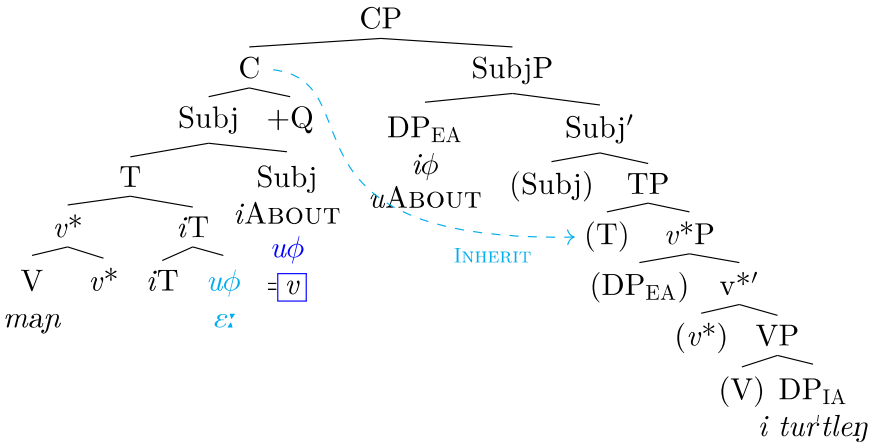
The data in (23) exemplifies a PRES tense form with an inverted enclitic SCL, as we saw previously in (1). The structure underlying data like this is essentially identical to the tree seen in both (20) and (22), except, as noted, with the addition of standard head-movement (T-to-C) of the inflected verb (24). Otherwise, the two uφs and Agree work exactly as above. In this context, the allomorph for the SCL.2PL morpheme is [v], a clear descendent of Latin VOS.[7] The relevant aspect of the context

[7] In line with Benincà's (1983) proposal for the diachrony of non-φ-related elements, discussed in the text above, it is in fact also possible to argue that the Bolognese proclitic allomorph for the SCL.2PL morpheme, [a], is a descendent of Latin VOS, with the only remaining segment being the modern descendent of the Latin vowel (cf. Cardinaletti & Repetti 2008:footnote 12). Contrary to accounts which make it impossible to treat the often-homophonous preverbal 1SG, 1PL, and 2PL elements as φ-related SCLs, the one here permits it, whatever the varying overt or null realizations are in particular grammars. Thus, while in Paduan and Donceto, the overt elements have evolved beyond φ-related SCLs, leaving only null proclitics present in these paradigm cells, Bolognese remains more conservative. Interestingly, the same Latin vowel occurs in the ancestor for the two Bolognese homophonous proclitic SCLs in 1SG and 1PL forms. Rohlfs (1968) shows that that ancestor (at least for the enclitic) is Latin EGO for both, with the singular extended to the plural at an early point in the history of Gallo-Italic. The same diachrony should extend to proclitics, despite other later obscuring changes, including reanalysis as not φ-related. In some grammars, distinct patterns of extension and other morphological changes, as proposed in Calabrese (2011), as well as

for this allomorph (which occurs in both PRES and FUT forms) is that it occurs in the coda of the tonic syllable. This complements the situation below, in IMPF (and PST and COND) forms, as we shall see.

(23) Bolognese, field work
maˈɲɛː=v i turˈtleŋ ?
eat.2PL= SCL.2PL the tortellini
'Are you eating the tortellini?'

(24) Proposed Interrogative, Present Tense



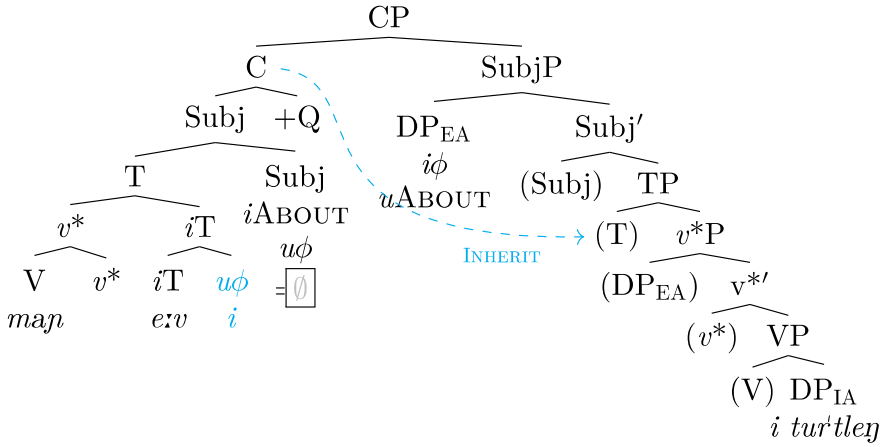
Moreover, and importantly, given this analysis, the observed T-to-C underlies the enclitic position of the SCL using only the standard logic of head-movement. Although this logic may be formalized in a variety of ways, we assume simple left-adjunction here. That is, just as V left-adjoins to *v** and the combination left-adjoins to T, so that the morpheme in V precedes the morpheme in T when both are overt as in (21) above or (25) below, so too does the form in T precede the SCL in Subj when T raises to Subj on its way to C. This analysis of SCLs does not need to add any mechanism for determining their ordering relative to the inflected verb, unlike previously available analyses, which stipulate distinct mechanisms for the ordering of proclitic vs. enclitic SCLs.

Finally, we consider IMPF data with enclitics, shown in (25). As seen in the tree in (26), once again, the essentials of the syntax below C, including the two distinct uφs and the effects of Agree, is the same as in all other data types. The effects of T-to-C movement on the morpheme ordering operate exactly as they did for the PRES data immediately above. Our proposal provides an entirely uniform syntactic analysis for all data types, with T-to-C movement the sole factor that distinguishes enclitics from proclitics.

(25) Bolognese, Field Work
ˈmaˈɲeːvi=Ø i turˈtleŋ
eat.IMPF.2PL=SCL.2PL the tortellini
'You were eating the tortellini.'

reanalysis, might occur. Finally, in the Bolognese proclitics, the preceding descendent consonants ([v] in 2PL and [j] in first person) were lost, but in the enclitics, they are retained.

(26) Proposed Interrogative, Imperfect Tense



In this example, the allomorph for the SCL.2PL is null, and, as noted above, the context is distinct from the one in (23-24). The enclitic here is positioned where it would, like above, be in a syllabic coda. However, in this IMPF data (and also in PST and COND data), the final syllable of the inflected 2PL verb is not tonic, as documented in the previous section. As a result, this sort of coda is distinct from the one in the final syllable of the inflected verb forms in PRES and FUT forms, where the allomorph is the overt [v]. The two sorts of codas are complementary, and this straightforwardly determines the difference in forms, without needing to invoke distinct syntactic structures for the two.[8]

5. Conclusion

This paper proposes an analysis of SCLs that relies entirely on a syntax that includes no *ad hoc* mechanisms. The primary assumption is that there is a secondary use of a generally available syntactic mechanism that is present in all grammars, namely *u*φ, which underpins the Case-licensing of argument DPs in all grammars as well as associated verbal agreement in grammars that have it. This second subject-related *u*φ entered early Gallo-Italic when the evolving descendants of Latin nominative tonic pronouns (then in the specifier of the functional head Subj that Rizzi (2015) independently motivated) were reanalyzed away from being some φ-feature bearing functional head in DP to being an additional *u*φ on Subj itself. Just as the standard *u*φ in the lower head T probes and Agrees with the highest argument DP, so too did this new *u*φ probe and Agree with that same DP, thus doubling the expression of agreement

[8] Bolognese has three allomorphs for the *u*φ in Subj valued as 2PL. Other grammars may also have three, though of course their particular realizations may vary. For example, in Ro Ferrarese, these appear to be [a] and the non-post-tonic zero (both as in Bolognese), though the post-tonic allomorph appears to be [f], rather than [v]. Other grammars may also have three distinct allomorphs with their own realizations, or they may have fewer or more realizations. For example, in the Donceto data from Cardinaletti & Repetti (2008), we have evidence for a null pre-verbal allomorph (under our analysis, and different from Bolognese), and for the [v] post-tonic allomorphs (like Bolognese), but there is no documented evidence for the other forms needed to determine whether or not there is a third.

with the subject. With this *u*φ now on Subj, SCL-doubling with a DP in the now vacated SpecSubj became possible, and began to be manifested.

The variable positioning of the SCLs relative to the inflected verb arises from the standard syntax in which they occur. A proclitic position arises when the two are in the separate heads mentioned above. An enclitic position arises when the lower head raises into the higher one by standard left-adjunction and carrying it on to their eventual position in C. This is standard head-to-head movement, of the T-to-C type. This syntax of this proposal is simple and explains the observed morpheme ordering without recourse to distinct mechanism for each ordering. Variability in the forms of proclitics vs. enclitics, as well as variability of form within each type, is left to the morphophonology.

This proposal also shifts the burden of explaining syncretisms and gaps in the paradigms of SCLs both within each grammar and across grammars from the syntactic to the morphophonological component of the grammar. This removes any need to posit additional distinct ad hoc functional categories for different specific SCLs, which is identical to the generally accepted lack of any need for additional distinct ad hoc functional categories for different specific tonic pronouns. The differences among both SCLs and tonic pronouns derive from the specific feature sets that each one contains and how those features are realized in the morphophonology. Although not addressed in detail here, this is also where the sorts of observable generalizations about cross-linguistic patterns concerning SCLs is captured.

Acknowledgments

I am extremely thankful to my consultants in Bologna for the great deal of time that they offered me, and for their patience and willingness to satisfy my curiosity, which was often inexplicable to them. I am also especially grateful to Roberto Serra and Daniele Vitale for extensive enthusiastic discussions of interconnecting issues in Bolognese grammar, and for introducing me to additional consultants and gathering additional data on my behalf. Finally, I thank two anonymous reviewers for inciteful commentary and probing questions that have helped improve this paper.

References

Ackema, Peter, & Ad Neeleman, 2018. *Features of Person*. MIT Press. doi.org/10.7551/mitpress/11145.001.0001

Benincà, Paola, 1983. Il Clitico ‘a’ nel Dialetto Padovano. In G. Pellegrini (ed.), Scritti Linguistici in onore di Giovan Battista Pellegrini, 25–32. Pisa: Pacini.

Benincà, Paola, 1997. Sentence Word Order. In M. Maiden & M. Parry, *The Dialects of Italy*, 123-130. New York: Routledge. doi.org/10.4324/9780203993880

Calabrese, Andrea, 2011. Investigations on markedness, syncretism, and zero exponence in morphology. *Morphology* 21: 283–325. doi.org/10.1007/s11525-010-9169-y

Canepari, Luciano & Daniele Vitali, 1995. Pronuncia e Grafia del Bolognese. *Rivista Italiana di Dialettologia* 19:119-164.

Cardinaletti, Anna, 2004. Towards a cartography of subject positions. In L. Rizzi (ed.), *The Structure of CP and IP*, 115-165. Oxford: Oxford University Press.

Cardinaletti, Anna & Lori Repetti, 2008. The Phonology and Syntax of Preverbal and Postverbal Subject Clitics in Northern Italian Dialects. *Linguistic Inquiry* 39: 523–563. doi.org/10.1162/ling.2008.39.4.523

Cardinaletti, Anna & Michal Starke, 1999. The Typology of Structural Deficiency: A Case Study in the Three Classes of Pronouns. In H. van Riemsdijk (ed.), *Clitics in the Languages of Europe*, 145-233. Berlin: Mouton de Gruyter. doi.org/10.1515/9783110804010.145

Ledgeway, Adam & Alessandra Lombardi, 2005. Verb movement, adverbs and clitic positions in Romance. *Probus* 17: 79–113. doi.org/10.1515/prbs.2005.17.1.79

Manzini, M. Luisa & Leonardo M. Savoia, 2005. *I dialetti italiani e romanci, Morfosintassi generativa*. Alessandria: Edizioni dell'Orso.

Paoli, Sandra, 2008. Interpolation structures and clitics in Triestino. In D. Bentley & A. Ledgeway (eds.), *Sui dialetti italoromanzi: saggi in onore di Nigel B. Vincent*, *The Italianist* 27, Vol. 3: 184–199. https://www.tandfonline.com/journals/yita20.

Rizzi, Luigi, 2000. *Comparative syntax and language acquisition*. New York: Routledge. doi.org/10.4324/9780203461785

Rizzi, Luigi, 2015. Notes on Labeling and Subject Positions. In E. Di Domenico, Hamann, C., & S. Matteini, *Structures, Strategies and Beyond: Studies in Honour of Adriana Belletti*, 17-46. Amsterdam: John Benjamins. doi.org/10.1075/la.223.02riz

Roberts, Ian, 2010. *Agreement and head movement: Clitics, incorporation, and defective goals*. Cambridge: MIT Press. doi.org/10.7551/mitpress/9780262014304.001.0001

Rohlfs, Gerhard, 1966-9. *Grammatica storica della lingua italiana e dei suoi dialetti*. Vol. 1-3. Turin: Einaudi.

Rubin, Edward J. & Aaron Kaplan, 2024. Segmental and Prosodic Influences on Bolognese Epenthesis. In J. Y. Kim, Miatto, V., Petrovic, A., & L. Repetti (eds.), *Epenthesis and beyond*. Berlin: Language Science Press. doi.org/10.5281/zenodo.14193821

Vanelli, Laura, 1987. I pronomi soggetto nei dialetti italiani settentrionali dal Medio Evo a oggi. *Medioevo Romanzo* 12: 173–211.

Vitali, Daniele, 2007. *Dscårret in bulgnais?* Bologna: Perdisa.

Zörner, Lotte, 1989. *Die Dialekte von Travo und Groppallo: diachrone und synchrone Studien zum Piacentinischen*. Vienna: Verlag der Österreichischen Akademie der Wissenschaften.