Romanian direct objects allow what at first seems like ordinary object “pro-drop”. The object can be expressed with a pronominal object clitic alone, as in (1), or with a nominal doubled by the clitic, as in (2). Our basic account of these data will follow traditional LFG analyses of pro-drop: the clitic has an optional [PRED ’pro’] feature, which means it is ambiguous between an agreement marker and an incorporated pronoun (Bresnan and Mchombo 1987, Fassi Fehri 1988, Bresnan et al. 2016: ch. 8). However, a closer inspection reveals that the Romanian data warrant a more elaborate analysis.

Our proposal draws on several different bodies of research. We rely on the traditional literature of grammaticalization of agreement marking (Givón and Li 1976, Mithun 1988), as well as recent literature stressing the importance of considering phonology/prosody and syntax separately when analyzing clitics (Lowe 2016, van Rijn 2016, Toivonen 2003: ch. 3). Our specific account is a straightforward adaptation of LFG pro-drop, coupled with the analysis of historical change developed in Coppock and Wechsler (2010). An important point we wish to highlight is that what may at first seem like odd quirks of Romanian in fact belong to a pattern that is typologically quite common. The data are also unsurprising given our theoretical assumptions. The formal machinery and architectural assumptions of LFG are well suited for modelling the kind of complex system found in present-day standard Romanian, and also for modelling the dialectal and diachronic variation of that system.

Romanian object pro-drop. The morphosyntax of Romanian direct objects involves differential object marking, as discussed by Ticio and Avram (2015), Tigău (2015), Hill (2013), von Heusinger and Chiriacescu (2013), and others. In essence, definite direct objects can be marked with pe as in (2), or unmarked for objecthood as in (3). We focus on the clitic doubling here and set aside the difference between (2) and (3) (there is disagreement on whether the differential object marking is best analyzed with reference to specificity, presupposition, or something else).

Only pe-marked objects can be doubled, in accordance with Kayne’s generalization (Anagnostopoulou 2005, Jaeggli 1982: 20), and only nouns, proper names and pronouns with human referents can be pe-marked. To capture these generalizations, we propose the lexical entry in (4a) for the agreement marking clitic. The agreement marker has no PRED feature, as the PRED is contributed by the nominal that it doubles. This particular clitic is third person singular masculine, but there are parallel clitics in other persons, numbers and genders as well. The agreement marker is specified with a positive human feature, as it never occurs with nominals with non-human referents. Finally, it has a constraining equation that demands the presence of an accusative CASE feature in its f-structure. The lexical entry does not itself contribute that feature, the feature must come from elsewhere. We treat pe as an accusative case marker which contributes just that feature. A noun without pe (e.g., baiatul in (3)) is not marked for accusative case and contributes no CASE feature; it can therefore not combine with the agreement marking clitic, and (3) cannot include a clitic double.

The lexical entry for the pronominal masculine singular clitic l- is given in (4b). Unlike the agreement marker in (4a), this clitic is a pronoun that itself has referential properties. Following standard LFG assumptions about pro-drop, this lexical entry has a PRED feature valued ‘pro’. The agreement marker and pronoun would typically be listed as a single entry with an optional PRED feature, to stress the connection between the entries. The Romanian facts warrant a more complex analysis than that: the agreement marker in (4a) is only compatible with humans, but the pronoun in (4b) has no animacy restriction. It can refer to inanimates, animals or humans. We also assume that it contributes its own CASE feature, instead of requiring such a feature to be provided elsewhere.

Grammaticalization. The analysis of Romanian pro-drop presented here involves two quite different lexical entries for the clitic in (1) and (2), even though the two are identical in form and similar in content (third person masculine singular). We argue that this is in fact not surprising when viewed from a diachronic perspective. Agreement marking is typically a result of a grammaticalization process along a cline: independent pronoun > weak pronoun > clitic pronoun > agreement affix > fused agreement
marker (Hopper and Traugott 1993). This type of change involves reanalysis: what is at one point analyzed as a pronoun is later analyzed as an agreement marker. In between, there can be a stage where the morpheme is analyzed as ambiguous between an agreement marker and a pronoun. At this stage there are in effect two lexical entries: one with a PRED feature and one without. The two lexical entries can change independently, as they have in Romanian. This has resulted in differences in the HUMAN and CASE feature, in addition to the PRED feature optionality. Coppock and Wechsler (2010) provide an in-depth account of how the grammaticalization process can affect features other than PRED (see also the papers in Butt and King (2001)).

The clitic is obligatory in examples like (2) (von Heusinger and Chiriacescu 2013). However, until recently, the clitic was optional (The Grammar of the Romanian Academy, 1963): (5) was permitted in earlier versions of the language. In the earlier variant, the clitic was a true pronoun that could be doubled by an adjunct pe-phrase. The clitic was then reanalyzed as an agreement marker when co-occurring with a pe-phrase, but remained a pronoun when standing alone. In other words, it was reanalyzed as having an optional PRED ‘pro’ feature. It is still possible to find dialects where the clitic is optional (von Heusinger and Chiriacescu 2013). Those dialects have not yet undergone the reanalysis, and the clitics cannot function as agreement markers in those varieties. That is, the PRED feature is still obligatory in those dialects, and doubling is only possible under certain discourse conditions (the object is a topic, co-indexed with the clitic).

The clitics in Romanian are true clitics and not bound affixes. This is true also when they function as agreement markers, even though agreement markers are typically affixes. van Rijn (2016) argues that this pattern is common as part of the grammaticalization process: loss of referentiality is not intrinsically tied to reduction in morpho-phonological form. In a modular framework like LFG, mismatches and unusual mappings between levels are expected to occur and readily modelled.

**TYPOLOGY.** If the assumptions laid out above are correct, we would expect the kind of lexical split we see in Romanian to occur in other languages as well. A typological survey indeed reveals several similar cases cross-linguistically; see, e.g., Hebrew subject agreement (Ritter 1995), Tundra Nenets possessors (Nikolaeva and Bárány, 2017), River Plate Spanish object agreement (Andrews 1990), and Finnish possessors (Toivonen 2000). We briefly review the latter two below.

Variations of clitic doubling occur across the Romance languages and dialects. In Andrews’s analysis of River Plate Spanish, the agreement marking clitic includes an \([\text{ANIMATE }+]\) feature that the pronominal clitic lacks. This is very similar to the Romanian analysis offered here. Estigarribia (2013) offers a different LFG analysis of the same Spanish variety, looking at a different angle of the data. He also proposes lexical entries where the agreement marker and the pronoun differ beyond the PRED feature: the agreement clitic is marked for specificity and the pronominal clitic is not.

In the nominal possession marking system in Finnish, the independent pronominal possessor is optional ((6); see Toivonen 2000). This is a typical instance of pro-drop which can be modelled with an optional PRED feature on the suffix -ni. This holds for first and second person singular and plural. However, third person possessive suffixes display more complex characteristics. Compare examples (7) and (8): the inclusion of the independent pronoun hän/en involves a distinction in meaning. In (7), the subject and the possessor cannot corefer, but in but in (8), the subject and the possessor must corefer. Furthermore, the possessive suffix can only agree with pronouns, and not with other types of nominals (9). In order to capture these facts, Toivonen (2000) posits the lexical entries in (10) for the agreement marking suffix and the pronominal suffix, respectively. The Finnish possessive suffixes are quite similar to the Romanian object clitics in that the agreement markers differ markedly from the pronouns.

**SUMMARY.** This paper revisits the traditional LFG analysis of pro-drop and applies it to Romanian direct object clitics. The clitics in Romanian display a lexical split remarkably similar to what is found in other agreement systems cross-linguistically. We argue that this kind of homophony of agreement markers and pronouns with differing characteristics is both common and unsurprising in light of (a) what we know about grammaticalization and the formal status of clitics; and (b) the standard LFG analysis of pro-drop, which already involves a lexical split in the sense that one entry includes a PRED feature and
the other does not. We further argue that the historical changes resulting in these lexical splits can be modelled with the system developed in Coppock and Wechsler (2010) for Uralic object agreement.

(1) L-am vazut.
    3.SG.MASC.ACC-have.I seen
    ‘I saw him/it.’

(2) L-am vazut pe baiat.
    3.SG.MASC.ACC-have.I seen ACC boy
    ‘I saw the boy.’

(3) Am have.I seen baiatul.
    ‘I saw the boy.’

(4) (a) l- (agreement)
    (↑ PERS) = 3
    (↑ NUM) = SG
    (↑ GEND) = MASC
    (↑ HUMAN) = +
    (↑ CASE ) = c ACC

    (b) l- (pronoun)
    (↑ PRED) = ‘pro’
    (↑ PERS) = 3
    (↑ NUM) = SG
    (↑ GEND) = MASC
    (↑ CASE ) = ACC

(5) Am vazut pe baiat. (archaic)
    have.I seen ACC boy.DEF
    ‘I saw the boy.’

(6) Jukka n¨akee sees yst¨av¨a-ni.
    ‘Jukka sees my friend.’

(7) Jukka n¨akee h¨anen yst¨av¨a-ns¨a.
    ‘Jukka sees his friend.’

(8) Jukka n¨akee yst¨av¨a-ns¨a.
    ‘Jukka sees his friend.’

(9) *Jukka n¨akee pojan yst¨av¨a-ns¨a.
    ‘Jukka sees his friend.’

(10) (a) -ns¨a (agreement)
    (↑ PERS) = 3
    (↑ PRONOMINAL) = +
    (↑ GEND) = c HUMAN

    (b) -ns¨a (pronoun)
    (↑ PRED) = ‘pro’
    (↑ PERS) = 3
    (↑ SUBJECT BINDING) = +

References


