Non-verbal predicate proforms: Category, case and φ-feature concord

Romance   Throughout the Romance languages, a definite clitic serving as a pro-predicate for an adjectival or indefinite nominal predicate is insensitive to the category and φ-features of its antecedent and local subject, resisting CONCORD, as shown by French M.SG *le* (1), Italian M.SG *lo* (2), and Catalan NEUT.SG *ho* (3); see (the references cited in) Déchaine & Wiltschko (2002), Espinal & Giusti (2023). This is remarkable in light of the fact that the clitic exhibits obligatory gender and number CONCORD with its antecedent when serving as an argument, and the fact that φ-feature concord between a [+N] predicate and its subject is otherwise robust in Romance. Catalan adds two further twists: the predicates in (3) can alternatively be replaced with the partitive/indefinite clitic *en*; and definite predicates are replaceable with the φ-concordial clitics *el/la/els/les*.

Uralic   While Meadow Mari uses the non-φ-marked indefinite proform *təgaj* 'so/such' as its pro-predicate in (4), Hungarian here cannot use *úgy* 'so' or *ugyanaz(ok)* 'same' (the latter irrespective of number marking) but instead exploits definite *az* 'it', which must inflect for plural number (the suffix *-k*) when it is case-concordial with its antecedent, as in (5), but strongly disprefers number concord when the pro-predicate differs from its antecedent in having dative case because it occurs in an agreeing infinitive (*lenni-ük* 'be-3PL'), as in (6). The number concord data in (5)–(6) match the behavior of non-proform predicates: replacing the pro-predicate with a full copy of the antedecent results in *gyönyörű-*(ek)* 'beautiful-PL' in (5) and *drágá-(^{??}k)-nak* in (6).

pro-D~pro-φ   French *l*-clitics can be used as pro-predicates (as in (1)) and as bound variables (*chaque homme* pense que tout le monde *l'aime* 'every man thinks that everyone loves him'). Déchaine & Wiltschko (2002) take this to argue that French *l*-clitics are not pro-D elements (which, for them, can only be arguments and cannot be bound) but pro-φ elements. The fact that the *l*-clitics double as definite articles then leads them to conclude that French definite articles are not D-heads but heads of φP — a conclusion they take to be confirmed by the expletive use of French articles: *Jean aime le vin* '(lit.) Jean loves the wine' is ambiguous between a referential reading of *le vin* and a generic one ('wine'). The Hungarian proform *az* (not discussed by Déchaine & Wiltschko) can likewise be used as a pro-predicate and as a bound variable; and like its French counterpart, the Hungarian definite article *a(z)* supports a generic reading in *János szereti a bort* 'János loves the wine'. This suggests a treatment of Hungarian *az* as a pro-φ element. But if French *le* and Hungarian *az* are both pro-φ elements, how to explain the fact that the two, when they are used as pro-predicates, behave differently with respect to φ-feature concord?

φ-Concord   Déchaine & Wiltschko's (2002) pro-φ analysis, while sound for *argument* clitics, is unsuitable for the uses of Romance definite clitics as proforms for adjectival or indefinite nominal *predicates*: these clitics are pro-N elements, showing no φ-inflection and representing the smallest possible size a non-verbal predicate can be in Romance (cf. the 'bare-NP' predicates in (1–3), (7)). Hungarian allows 'bare-NP' predicate nominals only if they are in the verbal modifier (VM) position (8a), not if the subject of predication is focused (8b). The subject of copular sentences featuring an unfocused pro-predicate is inevitably focused, hence the pro-predicate of (5) must be larger than a pro-NP: minimally a pro-φP, forcing number CONCORD with the subject.

Definiteness   Catalan can use its pro-φ clitics *el/la/els/les* as pro-predicates when their antecedent is definite: a case of definiteness concord. Espinal & Giusti (2023) also take concord to be behind the fact that in Catalan, when a left-dislocated predicate is introduced by *de* 'of', the pro-predicate resumptive clitic can be partitive/indefinite *en*: (9). But implicating concord for *en* is problematic: the use of *en* does not require the presence of *de*, nor does the use of *de* exclude *ho*. Déchaine & Wiltschko's (2002) treatment of French partitive/indefinite *en* as pro-N is in itself sufficient for Catalan (3), yielding free variation between the pro-N predicate proforms *ho* and *en*. French *en* and Italian *ne*, ungrammatical in (1) and (2), can only serve as pro-predicates within a complex noun phrase: their licensing requires nominal functional structure in their local domain.

Such* and *so   The forms *such* and *so* are not directly predicated of number-specified elements: they are predicates of degree. Since degree has no φ-features, (equivalents of) *such*/*so* cannot be φ-concordial. That the Mari pro-predicate *təgaj* 'so/such' in (4) resists plural *-vlak* is thus unsurprising; that *təgaj* is used instead of a nominal predicate proform is rooted in the highly restricted distribution of nominal reference-related markers across Uralic (see Simonenko 2014), recastable in Déchaine & Wiltschko's (2002) model with an appeal to their pro-D status. Hungarian *az*, being pro-φ, is the preferred pro-predicate choice over the more complex *ugyanaz* 'same', similar to the preference in standard English anaphora (except in the legal register) for *it* over *(the) same*.

Case matters   In Hungarian (6), the silent subject of the infinitive controls number AGREEMENT on the infinitive (*-ük*) analogously to number agreement in possessive nominals (*a pro*$_{3PL}$ *könyv-ük*$_{3PL}$ 'their book'), via a syntactic Spec–Head relationship. Den Dikken (1999) argues that *pro* inside the structural core has a dative-marked associate, occurring either in the left periphery of the nominal or infinitival clause or outside it (10). This associate is <u>not</u> in a syntactic AGREEMENT relation with anything inside the infinitival clause; however, being dative rather than (unmarked) nominative, it <u>is</u> the more case-specific choice for controller of case CONCORD with the predicate nominal. The dative associate, structured as in (11), has its φ-features embedded too low in its structure to make them directly accessible to subject–predicate concord. The result, for the predicate of the infinitival clause in (6), is obligatory DAT case concord but lack of φ-concord.

Conclusion   Déchaine & Wiltschko's (2002) pro-φ analysis of French *le* needs to be amended for Romance pro-predicate definite clitics (treated here as pro-N), but is fully adequate for all tokens of Hungarian *az*, whereas definite proforms are pro-D elsewhere in Uralic. The case and number concord facts in Hungarian (5–6) fall into place with an analysis of possessive nominals and inflected infinitives that features a peripheral dative KP that controls case concord but not φ-concord. The analysis affirms that a distinction is needed between AGREEMENT and CONCORD.

(1)   tu es belle/enseignante;   tes filles *le*/**la*/**les* seront aussi   (French)
       you are beautiful.F.SG/teacher.F.SG   your daughters CL.M.SG/*F.SG/*PL will.be also

(2)   tu sei bella/maestra;   *lo*/**la*/**le* saranno anche le tue figlie   (Italian)
       you are beautiful.F.SG/teacher.F.SG   CL.M.SG/*F.SG/*F.PL will.be also your daughters

(3)   tu ets bonica/mestra;   les teves filles també *ho*/**la*/**les*/*en* seran   (Catalan)
       you are beautiful.F.SG/teacher.F.SG   your daughters also CL.N.SG/*F.SG/*F.PL/IND will.be
       ALL: 'you$_{SG}$ are beautiful/a teacher; your daughters will be, too'

(4)   motor/okəktəʃo ulat;   üdəretvlakat *təgaj*(*-*vlak*) lijət   (Meadow Mari)
       beautiful/teacher you.are daughter.2SG.PL.ADD such(-*PL) be.3PL

(5)   gyönyörű/oktató vagy;   a lányaid is *az-ok*/**az*/**úgy*/**ugyanaz(ok)* lesznek   (Hungarian)
       beautiful/teacher you.are the daughter.PL.2SG also it-*(PL)/so/same(PL) will.be.3PL
       BOTH: 'you$_{SG}$ are beautiful/a teacher; your daughters will be, too'

(6)   az élelmiszerek nagyon drágák,   de nem kell *an-nak*/[??]*az-ok-nak* lenni-ük   (Hungarian)
       the groceries are very expensive but not need it-([??]PL)-DAT be-3PL
       'groceries are very expensive, but they don't need to be'

(7)   s'il était président   (8a)   ha [elnök]$_{VM}$ lenne   (8b)   ha Ő lenne [?]*(*az*) elnök
       if he were president   if president were   if (S)HE were the president

(9)   (*d'*)alegres, les criatures *ho*/*en* son   (Catalan)
       of happy.PL the children CL.DEF/IND are

(10)   *őnekik* (...) {a könyv-ük/nem kell menni-ük}   (Hungarian)
       they.DAT.PL the book-3PL/not need go-3PL   'their book / they don't need to go'

(11)   [$_{KP}$ K$_{[DAT]}$ (...) [$_{φP}$ φ [$_{NP}$ N]]]

References Déchaine & Wiltschko (2002): *LI*, https://doi.org/10.1162/002438902760168554 • Den Dikken (1999): in *Crossing Boundaries*; Benjamins • Espinal & Giusti (2023): *Linguistics*, https://doi.org/10.1515/ling-2022-0084