Pronouns and other proforms: Pro-predicates, expletives, demonstratives

---

- we have seen in previous sessions that pronouns can serve as **pro-predicates** — as in (1)–(4)

(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

(4) 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
all: 'you$_{SG}$ are beautiful/a teacher; your daughters will be, too'

→ while Hungarian cannot use *úgy* 'so' or *ugyanaz(ok)* 'same' (the latter irrespective of number marking) but instead exploits number-concordial definite *az* 'it', Meadow Mari uses the non-φ-marked indefinite proform *təgaj* 'so/such' as its pro-predicate in (5)

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

- *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

→ the fact that the Mari pro-predicate *təgaj* 'so/such' in (5) resists plural -*vlak* is unsurprising

→ the fact that *təgaj* 'so/such' 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*

- English *so* occurs as a predicate in (6a); it can also be used as a proform for predicates, in which case it is usually fronted to a position in the left periphery, as in (6b)

→ as a proform for predicates, *so* can 'stand in' for predicates of any category: (7)

(6) a. it is(n't) *so*
b. *so* it is (indeed)

(7) a. she is {beautiful/a teacher/in great shape}, and *so* are her daughters
b. she always gets up early, and {*so* do her daughters / and her daughters do *so*, too}

- but *so* is also used as a propositional anaphor, in cases such as (8) and (9), featuring epistemic verbs and *verba dicendi* — here *so* appears to be serving as an argument of the matrix verb

(8) A: did he do it?
B: I (don't) think *so*

(9) A: how do you know he did it?
B: because he said *so*

Q are we dealing with two different forms *so* — one a PREDICATE and the other an ARGUMENT?

- an integrated approach to *so* according to which it is uniformly an argument is clearly impossible: there can be no sense in analysing the *so* of (6) and (7) as an argument
- but a uniform analysis of *so* as a (pro)predicate *is* feasible, on the assumption that *so* in (8) and (9) is not *itself* a propositional anaphor

❶ *so* as a pro-predicate of a small clause in the complement of the verb (*think*, *say*)

→ its subject is a silent pronoun (*x*) representing the proposition in the antecedent

(10) $[_{VP}$ V=*think*/*say* $[_{RP}$ *x* $[_{R'}$ RELATOR $[_{Predicate}$ *so*]]]]]

→ (10) makes it relatively simple to account for an otherwise quite puzzling difference between English and Dutch
– English *I (don't) think so* expones the predicate (as *so*) and leaves the subject unpronounced (*x*)
– Dutch *ik denk het (niet)* 'I think it (not)' expones the subject (as *het*) and leaves the predicate unpronounced (SO)

BUT a hurdle for (10) is that *say* and *think* do not normally take small-clause complements

(11) a. *she said [it true]
b. %she thinks [it possible]

❷ *so* as a pro-predicate for the VP of the verb (*think*, *say*)

→ see the parallel between inversion with *so* and quotative inversion

(12) a. 'he's crazy,' thought Mary
b. 'he's crazy,' Mary thought

(13) a. 'he's crazy,' said Mary
b. 'he's crazy,' Mary said

(14) a. ... and Mary thought *so*, too
b. †... and *so* thought Mary

(15) a. ... and Mary said *so*, too
b. ... and *so* said Mary

→ quotative inversion should, in turn, be compared to locative inversion

(16) a. on this wall hung a picture of the president
b. down the hill rolled the baby carriage

(17) a. here comes Santa Claus
b. there goes my lunch

→ the propositional anaphor *so* is a pro-predicate for the VP of the *verbum dicendi* or epistemic verb, with V raising out of VP, at a minimum to *v*

(18) $[_{vP}\ \theta_{ext}\ [v\text{+}V\ [_{VP}\ \underline{\cancel{V}\ \text{PROPOSITION}}]]]$
$\Downarrow$
so

• Quirk *et al*. (1985): quotative inversion fails in compound tenses (* *'he's crazy,' was saying Mary*, * *'he's crazy,' would say Mary*, * *'what time is it?' had asked Mary*)

→ Collins & Branigan (1997:14): the same is true for inversion with *so* (**... and so had thought Mary*)

NB *go* can serve in lieu of quotative *say*: (19)

(19) a. so I go/went 'duh!'
b. so I go/went 'you're kidding, right?'

→ but neither quotative inversion nor *so*-replacement is attested with quotative *go*
– (20) is probably due to a register clash
– (21) (where (21b) contrasts with *as Maine goes, so goes the nation*) likely has its roots in grammar: *so* can only replace *subordinated* propositions (22); what follows quotative *go* is always a direct quotation, arguably not itself embedded within VP

(20) a. Mary went 'duh!/you're kidding, right?'
b. *'duh!/you're kidding, right?', went Mary

(21) a. *John went 'duh!/you're kidding, right?', and Mary went *so*, too
b. *John went 'duh!/you're kidding, right?', and *so* went Mary

(22) a. John said that he loved kimchi, and Mary said *so*, too
and *so* said Mary
b. John said 'I love kimchi', *and Mary said *so*, too
*and *so* said Mary

• just like pro-predicates oscillate between pronominal and non-pronominal forms, so do so-called **expletives** (a term used pretheoretically here; we will return to their true nature below)

→ in (23a), the filler of the structural subject position is a pronoun (*het* 'it'); (23b) is identical with (23a) except for the fact that here, the filler of the subject position is locative *er* 'there'

(23) a. *het* is gebleken dat hij gelogen heeft (Dutch)
it is turned.out that he lied has
b. *er* is gebleken dat hij gelogen heeft
there is turned.out that he lied has
both: 'it turned out that he lied'

NB 'expletive' *het* can control PRO; 'expletive' *er* cannot — see (24)
[the predicate of the adjunct in (24b) is different from that in (24a), to ensure that *er* is compatible with it: *lijken* 'seem' does not allow expletive *er* ({*het*/**er*} *lijkt plausibel dat hij gelogen heeft* 'there seems plausible that he lied'); but *beweerd zijn* 'have been alleged' does: *er is beweerd dat hij gelogen heeft* 'there has.been alleged that he lied']

(24) a. *het* is gebleken, zonder PRO aanvankelijk plausibel te lijken, dat hij gelogen heeft
it is turned.out without initially plausible to seem that he lied has
'it turned out, without initially seeming plausible, that he lied'
b. **er* is gebleken, zonder PRO ooit expliciet beweerd te zijn, dat hij gelogen heeft
there is turned.out without ever explicitly alleged to have.been that he lied has
intended: 'it turned out, without ever having been explicitly alleged, that he lied'

• **PRO** can never be an expletive (even in the absence of control)

(25) a. [for *there* to emerge a solution to this problem] requires a great deal more work
b. *[PRO to emerge a solution to this problem] requires a great deal more work

→ PRO cannot even be a proleptic pronoun in a clausal prolepsis construction — even though there are plausible ways of dealing with the proleptic pronoun as a non-expletive (and, as a matter of fact, an expletive approach to the proleptic pronoun is not particularly plausible, as we saw in the previous session)

(26) a. [for *it* to be certain that he lied] involves a heavy burden of proof
b. *[PRO to be certain that he lied] involves a heavy burden of proof

→ the ungrammaticality of (25b) and (26b) can be understood if *there* and *it* in the a–examples are underlying **predicates**: PRO can independently be shown not to be able to serve as a predicate, not even in predicate inversion constructions (in which the predicate is raised into an ungoverned structural subject position, creating a grammatical environment for PRO)

(27) a. the most promising candidate was Bill [before *it* was John]
b. *the most promising candidate was Bill [before PRO being John]

→ that PRO cannot be a pro-predicate *in situ* is easy enough to understand ('the PRO theorem')

→ in predicate inversion constructions, the predicate is raised into an ungoverned position (the structural subject position), which ought to make PRO legitimate

→ imagine, however, that 'the PRO theorem' is a *global* constraint on the distribution of PRO — holding not just at the end of the derivation but throughout: as soon as PRO is introduced into the structure, it must be prevented from being in a governed position

→ if PRO cannot be governed at *any* point in the derivation, the ungrammaticality of (27b) follows, and that of (25b) and (26b) can be connected to it (on a predicational approach to *there* and *it*)

→ for grammatical sentences such as (28a,b) it must then be assumed that the subject is base-generated in an ungoverned position (the structural subject position) — which is feasible on a 'delayed gratification' approach

(28) a. [PRO to love Donald] is a precondition for membership of the Republican Party
b. [PRO to be loved by Donald] is a precondition for membership of the Republican Party

---

• finally, we need to talk about **demonstratives** used as proforms

→ consider the alternation between the a– and b–examples in (29) and (30)
['CG'=common gender (M/F); 'NT'=neuter; *meisje* 'girl' is grammatically neuter; there is no gender distinction in the plural in Dutch]

(29) a. die jongen, ik mag {*hem*/?*die*/**dat*} niet (Dutch)
that boy I like him/DEM_{CG}/DEM_{NT} not
a′. die jongens, ik mag {*ze*/†*hen*/?*die*/**dat*} niet
those boys I like them/them/DEM_{CG}/DEM_{NT} not
b. die jongen, {*die*/**dat*/**hem*} mag ik niet
that boy DEM_{CG}/DEM_{NT}/him like I not
b′. die jongens, {*die*/**dat*/**hen*/***ze*} mag ik niet
those boys DEM_{CG}/DEM_{NT}/them/them like I not

(30) a. dat meisje, ik mag {*haar*/?*die*/?**dat*} niet
that girl I like her/DEM_{CG}/DEM_{NT} not
a′. die meisjes, ik mag {*ze*/†*hen*/?*die*/?**dat*} niet
those girls I like them/them/DEM_{CG}/DEM_{NT} not
b. dat meisje, {*dat*/%*die*/**haar*} mag ik niet
that girl DEM_{NT}/DEM_{CG}/her like I not
b′. die meisjes, {*die*/**dat*/**hen*/***ze*} mag ik niet
those girls DEM_{CG}/DEM_{NT}/them/them like I not

→ *die* and *dat*, when used in **contrastive left-dislocation** constructions, are usually referred to as '***d*-pronouns**' in the literature on Dutch

→ but this special terminology is unnecessary: '*d*-pronouns' are formally indistinguishable from demonstratives; demonstratives can be used as proforms for nominal phrases

• three things stand out in the data in (29) and (30)

(a) when the associate of the left-dislocated constituent is spelled out **clause-internally**, a demonstrative proform is never the preferred option; but *die* is relatively acceptable while *dat* is impossible — regardless of the grammatical gender of the left-dislocate (thus, though *meisje* is grammatically neuter, *die* is much better than *dat* in (30a))

(b) when the associate of the left-dislocated constituent is spelled out in the **left periphery**, a demonstrative proform is the only option; for many speakers, the demonstrative in the left periphery exhibits grammatical-gender **concord** with the left-dislocated constituent

(c) personal pronouns exhibit natural-gender **concord** with the left-dislocated constituent

NB the variation regarding the grammatical-gender form of the demonstrative in (30b) matches the variation on this point in relative clauses, likewise introduced by a demonstrative

(31) a. de jongen {*die*/**dat*} ik graag mag
the boy DEM_{CG}/DEM_{NT} I gladly like
b. het meisje {*dat*/%*die*} ik graag mag
that girl DEM_{NT}/DEM_{CG} I gladly like
c. de jongens/meisjes {*die*/**dat*} ik graag mag
the boys/girls DEM_{CG}/DEM_{NT} I gladly like
'the boy(s)/girl(s) that I like a lot'

[note that in English the demonstrative-like element introducing relative clauses never shows φ-feature concord with the head (cf. Dutch (31c) with *the boys* {*that*/**those*} *I like a lot are Tom, Dick and Harry*) — in English, the demonstrative *that* has grammaticalised as a complementiser in relative clauses and other subordinate clauses alike; in Dutch relative clauses, the demonstrative is a phrasal element, not a complementiser, and it must show gender concord with the 'head'; variation on (31b) revolves around the question of whether concord involves grammatical or natural gender]

• **HYPOTHESES**

(i) the demonstrative and the left-dislocated constituent are base-generated separately in the syntax of the a–examples

(ii) the demonstrative and the left-dislocated constituent are base-generated as a constituent, entertaining a Spec–Head relationship, in the syntax of the b–examples (cf. Grohmann)

(32) $[_{DP}\ [jongen/meisje]_i\ [_{D'}\ DEM\ (t_i)]]$

• the syntax in (32) is a 'big DP' similar to the one proposed in some of the literature for **clitic doubling** constructions

→ analogously to φ-agreement in clitic doubling cases, (32) ensures φ-feature matching between the demonstrative and the left-dislocated constituent in the same way as in relative clauses [(32) which leaves open the size of the constituent projected by *jongen*/*meisje* 'boy/girl' as well as the question of whether this constituent ends up in SpecDP via movement or base-generation: relatives and left-dislocation constructions may differ precisely on these points]

- the demonstrative and the left-dislocated constituent are split in the course of the derivation, as a result of movement of the left-dislocated constituent into a higher position in the left periphery

(33) $[_{XP}\ [\textit{jongen/meisje}]_i\ [_{X'}\ X\ [_{YP}\ [_{DP}\ t_i\ [_{D'}\ DEM\ (t_i)]]]\ [_{Y'}\ Y\ ...\ ]]]]$

[the exact nature of 'X' and 'Y' is immaterial for present purposes — both are heads in the 'topic field' of the extended clause; the only thing that matters when it comes to deriving the strings in the b–examples in (29) and (30) is that 'X' remains empty while 'Y' receives the finite verb]

- by hypothesis, the 'big DP' in (32) can only have a **demonstrative** as its head — since Dutch is not a 'clitic doubling' language, it does not allow pronouns to team up with an associate in SpecDP

→ this accounts for the ungrammaticality of the b–examples in (29) and (30) with a personal pronoun

- by hypothesis, the 'big DP' in (32) is only allowed to surface in **derived** positions — for clitic doubling, too, this is a common assumption in the literature (though recall Paparounas & Salzmann 2023)

→ the obligatory displacement of the 'big DP' accounts for the fact that grammatically φ-concordial demonstratives are impossible in clause-internal position

- when a demonstrative or personal pronoun shows up in clause-internal position as the associate of a left-dislocated constituent, that demonstrative/pronoun is base-generated independently of the left-dislocated constituent (hypothesis ***(i)***) and serves as a **resumptive**

→ resumptive pronouns are not in a grammatical φ-concord relationship with their antecedents; they do, however, show φ-concord for natural gender, consistently yielding non-neuter *die* (never neuter *dat*) for [+HUMAN] left-dislocates

- for further discussion of (the literature on) contrastive left-dislocation, and for a presentation and discussion of the facts from Hungarian in this connection, see Den Dikken & Surányi (2017), 'Contrasting contrastive left dislocation explications', *Linguistic Inquiry* 48, 543–584

NB the topic originally announced for the session on 8 May ('Pronouns and referential dependencies: Binding and coreference') will **NOT** be covered in a session of its own — see pp. 9–10 of handout 1 for relevant discussion, which will have to suffice because of lack of time

→ instead of a session on referential dependencies, the class on 8 May will, like the one on 15 May, be devoted to student presentations of research proposals