Pronouns and obligatory displacement (II): Cliticisation

- in the previous session, we saw that WEAK pronouns have a syntactic distribution different from that of STRONG pronouns and full DPs — one that implicates **phrasal** movement to a specifier position outside or on the edge of the verbal core

- in this session, we will look at the syntactic distribution of CLITIC pronouns

→ both Cardinaletti & Starke (1999) and Déchaine & Wiltschko (2002) take clitics to be elements that must be hosted by a **head** — i.e., the end product of clitic placement is a head-adjunction configuration

(1) $[_{XP}\ [_{X}\ X{+}CL]\ [_{YP}\ \ldots\ ]]$

→ but although there is a broad consensus on clitics being hosted by (i.e., adjoined to) heads, there is much debate in the literature on how (1) comes about:

(i) (1) is base-generated; the clitic is coindexed with a silent element (*pro*) in a θ-position (for argument clitics) or in the position of the predicate (for pro-predicate clitics); movement of the silent element to SpecXP may be involved, for licensing purposes

(ii) (1) results from movement; the clitic originates in a θ-position (for argument clitics) or in the position of the predicate (for pro-predicate clitics) and moves directly to X

(iii) (1) results from movement; the clitic originates in a θ-position (for argument clitics) or in the position of the predicate (for pro-predicate clitics) and moves in a two-step process: *(a)* phrasal movement to SpecYP followed by *(b)* head-movement of the clitic to X

(iv) (1) results from movement; the clitic originates as a subpart of a complex constituent ('big DP') in a θ- or predicate position, and subextracts from its container (in one of the two ways sketched in *(ii)* and *(iii)*)

- ***(i)*** is the null hypothesis regarding the *placement* of the clitic

→ it superficially likens clitics to **agreement markers**, which are not arguments or predicates themselves

→ this raises the non-trivial concern that it is well known that clitics and agreement markers behave differently in a number of non-trivial ways:

– agreement markers are in a fixed position relative to their hosts, immutable as a result of syntactic processes (such as I-to-C movement); clitics can change places with their hosts under certain circumstances

(2) a. je *le* fais [**proclisis**] (French)
I CL:M.SG do

b. fais-*le*! [**enclisis**]
do-CL:M.SG

– clitics can be attached to non-finite verb forms that cannot host agreement markers

(3) je veux [*le* faire] (French)
I want CL:M.SG do.INF

– clitics can sometimes be variably placed, either on the selecting non-finite verb or on a higher verb ('clitic climbing'); variable placement of agreement markers does not occur

(4) a. voglio far*lo* (Italian)
want.1SG do.INF.CL:M.SG
b. *lo* voglio fare **lo* [vol-faccio]
CL:M.SG want.1SG do.INF CL:M.SG want-do.1SG

– in their phonological form, clitics are generally insensitive to the tense of their host; agreement markers often do show tense sensitivity

(5) a. je *le* fais (French)
I CL:M.SG do.PRES
b. je *le* faisais
I CL:M.SG do.IMPF
c. je *le* ferai
I CL:M.SG do.FUT
d. je veux *le* faire
I want CL:M.SG do.INF
e. en *le* faisant
in CL:M.SG do.PRESPTCP

(6) a. csinál-*ok* csinál-*sz* (Hungarian)
do(PRES)-1SG.INDEF do(PRES)-2SG.INDEF
b. csinál-t-*am* csinál-t-*ál*
do-PAST-1SG do-PAST-2SG.INDEF
c. csinál-n-*om* csinál-n-*od*
do-INF-1SG do-INF-2SG

→ a general direct assimilation of clitics to agreement markers is not feasible (even though the relationship between clitics and agreement markers is often close, and the treatment of φ-morphology as a clitic or agreement marker is frequently subject to intense debate)

• ***(ii)*** is the opposite extreme, assimilating clitics to their non-clitic counterparts in their underlying syntax, and attributing to clitics only one property that sets them apart from their non-clitic counterparts: their need to undergo head-movement to a head that can host them

(7) $[_{XP}\ [_{X}\ X+CL_i]\ [_{YP}\ \ldots\ [_{ZP}\ Z=t_i]\ \ldots\ ]]$

→ we have already seen that clitics can occasionally 'climb' to a verb that does not select the phrase in which the clitic is supposed to originate — for ***(ii)***, this introduces a complication: in (4b), the clitic must **skip** one or more intermediate heads in the process of head-moving to its host

→ well-established cases of head-movement are robustly local, resisting such skipping (see (8)) — the **Head Movement Constraint** (reducible to the **Minimal Link Condition**)
[in the literature on the languages of the Balkans, there is discussion of 'long head-movement'; but the facts involved are amenable to other treatments, not violating the HMC/MLC]

(8) a. he could have done it
b. could he have done it?
c. *have could he done it?

• ***(iii)*** was introduced with an eye towards factoring 'head skipping' out of the syntax of cliticisation, thanks to its two-step movement process

(9) $[_{XP}\ [_{X}\ X{+}CL_i]\ [_{YP}\ [_{ZP}\ Z{=}t_i]_k\ [_{Y'}\ Y \ldots\ t_k\ \ldots\ ]]]$

→ for nominal argument clitics, movement from ZP to SpecYP can be linked to **object shift**; once ZP has arrived in SpecYP, local to X, the clitic can head-move to X without needing to skip any intervening head along the way

→ note, however, that while (9) avoids a run-in with the HMC/MLC, it comes with its own penalty: the clitic has to extract from a derived specifier — this is something that established cases of head-movement (such as noun incorporation) generally do not tolerate: indeed, head-movement out of specifiers is often proscribed even for underived specifiers (Baker 1988 on the ban on N-incorporation involving the subject of unergative verbs: [*child-birth*], [*plane-crash*], [*brain-freeze*], [*temperature-drop*/*rise*], [*snow-fall*/*melt*] (all ergative) ~ *[*child-cry*], *[*clock-tick(ing)*], *[*horse-stumble*] (cf. [*horse-jump*], where *horse* is the object of *jump*), *[*man-run*] (cf. [*dog-run*], where *dog* is the object of *run*))

→ a more sophisticated argument against the two-step derivation comes from the French data in (10) (based on Kayne 1975, *French syntax*)

(10) a. je ne veux <u>rien</u> que tu *lui* dises (French)
I NEG want nothing that you him/her tell
'I don't want you to tell him/her anything'
a′. *je ne *lui* veux <u>rien</u> que tu dises
I NEG him/her want nothing that you tell

b. il faut <u>tout</u> que je *leur* enlève
it is.necessary everything that I them take.off
'I have to take everything off them (mother and children)'
b′. *il *leur* faut <u>tout</u> que j'enlève
it them is.necessary everything that I take.off

  - Kayne points out that for a subset of speakers of French, *rien* 'nothing' and *tout* 'everything' can be extracted out of a subjunctive clause via a process that can be likened to object shift
  - but in the very same sentences in which long-distance movement of *rien*/*tout* takes place, clitics belonging to the subordinate clause still cannot reach the matrix clause
  - if an object shift-like step precedes cliticisation, it remains unclear why upstairs clitic placement should be impossible in environments in which *rien*/*tout* can shift into the matrix clause

- ***(iv)***, the 'big DP' analysis, faces the same problems as ***(ii)*** if the clitic is assumed to head-move straight from the DP to its host, and the same problems as ***(iii)*** if the clitic is assumed to head-move only after the 'big DP' has undergone movement to a position local to the host

- the 'safest' approach to cliticisation remains one along the lines of ***(i)*** — i.e., non-movement of the clitic

→ but the clitic must then be taken to be in a different structural position relative to the host from the position occupied by agreement morphology, in order not to fall into the trap of inadvertently identifying clitics and agreement markers

→ one possibility that has been explored is that agreement morphology is attached *below* X^0 (say, at X^{-1}) while clitics are attached *to* X^0

(11) $[_{X0}\ [_{X0}\ [_{X-1}\ X{+}AGR]]\ CL]$

→ this would also help explain the fact that clitics are typically attached *outside* agreement morphology — BUT there are exceptions to this: **mesoclisis** in European Portuguese, for instance

(12) tu és professora; as tuas filhas sê-*lo*-ão também (European Portuguese)
you are teacher.F.SG the your daughters be-CL:M.SG-AGR:3PL too

→ mesoclisis patterns of the type in (12) suggest that both clitics and agreement morphology are syntactically relatively independent of the verb

→ this prompts an investigation of the possibility that clitics are not attached to lexical heads but to **functional heads** instead

- Kayne (1991) points out that, though clitics are usually adjacent to their verbal hosts, (literary) French occasionally allows a proclitic and its infinitival host to be separated by adverbial material — and this adverbial material can itself be clearly phrasal (as in (13b))

(13) a. pour *le* <u>bien</u> faire (French)
for CL:M.SG well do.INF
b. *en* <u>fort bien</u> parler
CL:PART strong well talk.INF

→ these facts indicate that clitics are not attached to the lexical verb — at least, not in the cases in (13)

→ this leaves open in principle the possibility that in cases in which the clitic and the host *are* adjacent, adjunction to V is involved

→ but a **unified analysis** of *all* clitic placement facts (in French, including (13), and beyond) will require abandoning the hypothesis that the clitic takes the lexical verb as its host

Q do clitics attach variably to different hosts, or to the left or to the right of their host, or is the locus and side of adjunction universally the same?

• Kayne (1991) argues that an **integrated approach** to clitic placement in the Romance languages is feasible only if clitics uniformly attach on the same side of the same host

(14) a. je *le* fais (French)
I CL:M.SG do.1SG
b. je veux [*le* faire]
I want CL:M.SG do.INF

(15) a. *lo* faccio (Italian)
CL:M.SG do.1SG
b. voglio [far*lo*]
want.1SG do.INF.CL:M.SG

→ it is known independently of the clitic-placement facts that, while verbs are uniformly raised out of the verbal domain in finite clauses in Romance, French and Italian differ with respect to verb movement in infinitives (see Pollock 1989 on French, Belletti 1990 and Kayne 1991 on Italian)

(16) ne ⟨*faire⟩ jamais ⟨faire⟩ une erreur (French)
NEG make.INF never make.INF an error

(17) non ⟨fare⟩ mai ⟨fare⟩ un errore (Italian)
NEG make.INF never make.INF an error

→ the difference between (14b) and (15b) should be related to the difference between (16) and (17), and should ideally fall out automatically as a function of the distribution of verb movement out of the verbal domain

(18) NEG X *jamais*/*mai* (...) [$_{VP}$ V]

→ the infinitive can reach X in Italian but not in French

→ assuming that there is a place between X and VP where clitics are located, we can derive the difference between (14b) and (15b) straightforwardly from the extent of infinitive movement
– in Italian, V.INF raises to X, past CL, hence surfaces to the left of clitics
– in French, V.INF does not raise to X, hence surfaces to the right of clitics

(19) X CL (...) [$_{VP}$ V]

NB CL in (19) could either be in a head distinct from X (e.g., Sportiche's 'CliticVoice') or be hosted by X

→ if X is the host of CL, the question of whether CL is left- or right-adjoined to it becomes relevant

→ right-adjunction might seem the most straightforward; but Kayne (1994) presents a possible analysis wherein CL is left-adjoined to X and the raised infinitive, in turn, left-adjoins to CL

- (19) also directly accounts for the possibility of adverbial insertion between the clitic and the infinitive, as in (13): this is the output obtained via not raising the infinitive out of VP

- the full account of verb and clitic placement in Romance must be considerably more complex than this, in light of the French facts in (20), which we will not attempt to analyse in detail here (because doing so would lead us too far afield, beyond the morphosyntax of pronouns)

  [this is a legitimate (albeit challenging) topic for a research proposal for this seminar]

(20) a. ne pas *l'*embrasser souvent (French)
NEG not CL:SG kiss.INF often
b. ne pas souvent *l'*embrasser
c. *ne pas *la* souvent embrasser

- the message that the discussion of Romance verb and clitic placement brings us is that clitics (at least in these languages) are uniformly hosted by a **functional** head
- though the generative syntax of cliticisation has primarily been developed on the basis of data from the Romance languages, there is a whole world of clitics outside Romance, which we will not have a chance to look at in this seminar
- the comparative morphosyntax of cliticisation is an important topic for research — you could in principle pick any clitic system that interests you and develop its morphosyntax against the background of what has been argued for Romance
- **Hungarian** is a language for which it has not been customary to assume that it has clitics — but in my own work, I have argued that Hungarian does have (at least) two *bona fide* clitics
  - the marker *-j*/*i* (part of the 'definiteness agreement marker' in present tense clauses) is traced back to an object clitic (**se*) in Proto-Uralic, and can be analysed in the same terms in present-day Hungarian
  - the *-lak*/*lek* form is a composite consisting of the first person singular subject-agreement morpheme *-k*, an epenthetic vowel, and a second person object clitic *-l*

  [see Den Dikken (2005) ('When Hungarians agree (to disagree): The fine art of 'Phi' and 'Art'' (keynote talk, ICSH7) and Den Dikken (2018) ('An integrated analysis of Hungarian nominal and verbal inflection', in Bartos, Den Dikken, Bánréti & Váradi (eds), *Boundaries crossed, at the crossroads of morphosyntax, phonology, pragmatics and semantics*, 147–62)]