**Language Optionality and Variation**
Festschrift in Honor of Anna Cardinaletti
edited by Chiara Branchini, Gianluca Lebani,
Lara Mantovan, Francesca Volpato

# Clitics Versus Modularity

Michal Starke
MUNI Masaryk University, Czech Republic

**Abstract** The core syntactic properties of clitics seem to universally covary with the core phonological properties of clitics and with the core semantic properties of clitics. In a truly modular world, that should be impossible. If the mind is in fact modular, what are those properties of clitics telling us? What used to be considered 'semantic' properties of clitics (such as animacy) are now widely considered syntactic, resolving the tension on the semantics side. This contribution explores the idea that the tension on the phonological side is resolved identically: the so-called phonological properties of clitics are in fact syntactic properties. This makes all properties of clitics syntactic, trivially satisfying modularity. It also entails that the syntactic functional sequence grows to integrate projection such as WordStressP, just like it grew to integrate projections such as AnimacyP.

**Keywords** Clitics. Modularity. Syntax. Stress. Focus. Universals.

**Summary** 1 Introduction. – 2 Clitics Versus Modularity. – 3 The Evolution of Semantics Shows a Way Out. – 4 Phonology: Rhythmic Versus Melodic. – 5 Conclusion.

## 1 Introduction

Cardinaletti and Starke (1999), henceforth C&S, emphasize the cross-modular nature of (syntactic) clitics: The peculiar syntactic behaviour of clitics has generated a rich literature, but so has their phonological behaviour, and C&S document that their semantic behaviour is no less unorthodox. What has been less emphasized is the challenge this creates for modularity. The syntactic, phonological and semantic behaviours of clitics correlate systematically across languages: in language after language, the same set of syntactic behaviours (of

Edizioni Ca'Foscari

**LiVVaL 9**
e-ISSN 2974-6574 | ISSN 2974-6981
ISBN [ebook] 979-12-5742-086-4 | ISBN [print] 979-12-5742-121-2

**Peer review | Open access**
Submitted 2025-11-27 | Accepted 2026-01-16 | Published 2026-11-17
© 2026 Starke | CC BY 4.0
**DOI 10.30687/979-12-5742-086-4/013**

clitics) correlates with the same set of phonological behaviours, which in turn correlates with the same set of semantic behaviours. Language thus seems to have a way to enforce that a given property inside syntax *systematically* corresponds to a given property inside phonology (which *systematically* corresponds to a given property inside semantics). That should not exist, in a modular world.[1]

In a truly modular world, what happens inside one module is inaccessible to what happens inside another module. This is a core property of modules, which Fodor (1983) calls 'encapsulation'. The only way to relate two distinct modules is through a 'dictionary' (lexicon) translating the output of one module into the language of the next module. The content of such translations, the lexicon, is however notoriously language-specific while the essence of the generalisation about clitics is language-invariant. There is thus no way to capture the facts about clitics in a strictly modular world (with language-specific lexica).[2,3]

What do we do with the facts of clitics then? It turns out that if we bite the bullet and pursue the modularity logic to its conclusion, we land in a place that may feel uncomfortable at first, but quickly becomes intriguing and hopefully insightful: some parts of phonology are in fact syntax; there may be functional projections of the type 'wsP', 'wordStressPhrase'. Let's get there step by step.

## 2 Clitics Versus Modularity

Perhaps the most discussed property of syntactic clitics is that they occur in different positions than the non-clitic members of their word class. In (1), the noun *Jean* and the strong pronoun *lui* have access to the same position, but the clitic *le* cannot occur there:

| | | | | | | |
|---|---|---|---|---|---|---|
| (1) | C'est | { Jean / lui / *le } | que | tu | vois |
| | *It's* | *J/him* | *that* | *you* | *see* |

**1** The work on this chapter was supported by Czech Science Foundation (GAČR) grant no. GF23-04856K.

**2** There is a second fact preventing generalisations spanning the content of two modules, what Fodor (1983) calls 'domain specificity'. I will elide this important issue to keep the discussion focused, but see footnote 5.

**3** The McGurk effect, whereby the processing of colors interferes with the verbal processing of color words can be interpreted as a challenge to encapsulation. This is however only true if it can be shown that the processing of colors interferes with the *inner* working of the verbal processing module. What the McGurk effect shows however is not that: it shows that the output of verbal processing is affected, which can also be due to the difficulty of higher level integration of the information when color processing tells you something (apparently) incompatible with what verbal processing is telling you.

Conversely, the clitic has access to a position unavailable to either the noun or the strong pronoun:

(2) Tu { *Jean / *lui / le } vois

*It's J/him see*

*"You see John/him"*

This syntactic distinction correlates with a seemingly arbitrary difference inside semantics: strong (personal, non-demonstrative) pronouns must be +human/animate, whereas clitics don't need to. As a result, the pronoun in the previous example can refer to either an animate or an inanimate entity:

(3) a. **Jean** tu **le** vois
*J, you him see*
*"You see John"*

b. ***ce*** ***livre***, *tu* ***le*** vois
*this book, you him* see
*"You see this book"*

The initial clefted example on the other hand cannot refer to the book:

(4) a. **Jean** c'est **lui** que j'ai vu hier soir
*J, it's him that I've seen yesterday evening*

b. ****ce livre*** *c'est* ***lui*** que j'ai vu hier soir
*this book it's him that I've seen yesterday evening*

This is also true of what looks like one and the same pronoun: *lui* in its clitic distribution allows inanimate referents, but disallows them in its strong pronoun distribution, a classical case of syncretism.

(5) a. ***ce livre***, je **lui** ai fait une place dans l'armoire
*this book, I him have made a space in the cupboard*
*"I made space in the cupboard for this book"*

b. **** ce livre***, *c'est* ***lui*** *que* j'ai vu hier soir
*this book, it's him that I've seen yesterday evening*

What happens inside syntax thus correlates with what happens inside interpretation. This correlation goes far beyond clefting and animacy. C&S document a range of other semantic distinctions between clitics and strong pronouns, and the syntax of clitics goes well beyond their special positioning. The question is how to capture these correlations between syntax and interpretation.

The cross-modular issue becomes perhaps more salient when the correlation is extended to phonology. The most celebrated property of clitics on the phonological side is their lack of word stress.[4] The clitic use of *lui* in French for instance does not induce any accentuation in unmarked spoken language (indicated below by the apostrophe linking the contraction of 'que' and 'son'), whereas the strong pronoun use of apparently the same *lui* does induce a pitch accent. Let's start with an unaccented baseline such as a possessive *son* in the initial position of a coordination, (6a).

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (6) | a. | … qu' | son | frère | et | sa | soeur |
| | | *… that* | *his* | *brother* | *and* | *his* | *sister* |
| | b. | … que | lui, | et | sa | soeur | |
| | | *… that* | *him,* | *and* | *his* | *sister* | |

When that initial position is substituted by a pronoun, the pronoun surfaces as a strong pronoun, since coordination disallows clitic pronouns (and more generally deficient pronouns). That strong pronoun comes with a slight accentuation to my ear, indicated by the underline on *lui* in (6b). The clitic use of the same-looking *lui* however doesn't induce any accentuation:

| | | | | |
|---|---|---|---|---|
| (7) | Jean | lui | parlera | demain |
| | *John* | *him* | *will.speak* | *tomorrow* |
| | *"John will speak to him tomorrow"* | | | |

It is thus specifically a property of the clitic nature of *lui* to lose word-stress. While the above facts are particular to French, the general description is recurrent across languages. The Slovak 3.sg.acc clitic *ho* is for instance also stressless and uncoordinable while its strong pronoun counterpart *jeho* has both word stress and the capacity to enter coordination. Clitichood thus has phonological behaviour acting in lockstep with syntactic behaviour – and with semantic behaviour. How can the inner working of phonology know about the

**4** C&S document that clitics are in fact capable of carrying phrasal stress, and even a certain subtype of contrastive stress. But not word stress.

inner working of semantics, what's more in an apparently language-invariant manner?[5]

## 3 The Evolution of Semantics Shows a Way Out

Once upon a time, theories of natural language syntax were entirely made out of arbitrary, non-interpreted symbols, such as S, NP, etc. designed to capture word-order.[6] In that world, the fact that the ability of clitics to coordinate or appear in particular positions correlates with their ability to refer to a *[-animate]* entity is a modularity scandal. Two modules see the insides of each other.

Fairy tales are however not known to be particularly realistic, and neither is this 'once upon a time' story. Syntax has never been free of semantically interpreted elements. Including within schools of thought which emphasize the autonomy of syntax, such as generative grammar. The semantically interpreted notion of tense has for instance been a key player inside syntax right from the start of the generative tradition, Chomsky (1957). The presence of semantically interpreted properties inside syntactic computation only grew from there. In the sixties and seventies, notions such as negation and questions took a more central and ever growing role in capturing word-order generalisations; and that trend accelerated sharply since the late eighties, incorporating notions such as focus, aspect, definiteness, specificity, plurality, etc. as generalisations about syntax became more fine-grained.

In the real (contemporary) world then, animacy is not clearly a notion outside of syntax. In fact, the word-order facts of various languages suggest that it very much is a syntactic notion, on the same level as negation, tense, definiteness, questions, aspect and many others. This was already clear enough back in the nineties that C&S didn't hesitate to assign the animacy/human feature to a particular functional projection of the noun phrase, within syntax. Very much the same story holds of other semantically interpreted properties of

**5** One may object that lexical stress is not about 'what happens inside the phonological module', but is rather about static lexical representations. That may or may not be true, given that what looks like the same lexical entry gains or loses word stress depending on its use as a clitic or a strong pronoun. Supposing it is true however simply shifts the issue from encapsulation to domain specificity, leaving the modularity puzzle intact – see footnote 2. Either way, these facts shouldn't exist if modularity is correct, the lexicon is language-specific, and phonology is a distinct module.

**6** One could claim that this is the very definition of syntax, as opposed to semantics. The term 'syntax' is however used here as the name of the mechanism responsible for word order in humans, making it an empirical question whether that mechanism is distinct or not from (some of the) mechanisms responsible for interpretation in humans.

the clitic/strong distinction such as their generic and expletive (non-referential) uses, focus, etc.

In that world, the modularity problem between the syntactic and semantic properties of clitics doesn't arise: all of the 'semantic' properties under discussion turn out to live inside the syntax module.

Does that mean that all of what we think of as semantics lives inside syntax? It turns out that 'semantics' splits into two, along traditional lines. On one side we find semantic properties that are recurrently relevant to word order generalisations – those largely correspond to formal, compositional semantics (and involve notions such as animacy, tense, quantification, specificity, modality, etc). On the other side, a large amount of semantic distinctions are never relevant to syntactic generalisations – those correspond to what is sometimes called lexical (or conceptual) semantics (and involve notions such as tables, carburettors, compassion or chickpeas). The historical change is that formal, structural, semantics notions have been gradually absorbed by syntax. Whereas conceptual semantics notions stayed out.

## 4 Phonology: Rhythmic Versus Melodic

What about the correlation of phonological word-stress with animacy, coordination, movement, etc? That looks like a deeply phonological property working hand in hand with non-phonological properties.

Let's start by noticing that the phonologically relevant properties of clitics are always about "structural" phonology, never about melodic phonology: they are always about stress (relative prominence of different chunks) and maybe syllabification (if clitics are always monosyllabic). There is no universal about the fricative, or labial, or nasal properties of clitics. The phonological properties relevant to clitics are also well-known to regularly interact with syntax: there are many generalisations about syntactic domains affecting stress assignment (or syllabification). On the other hand, the melodic properties of phonology, which are irrelevant to clitics, also never interact with syntax – there are famously no generalisations linking syntactic behaviour to being a fricative or labial, or nasal (Zwicky, Pullum 1986). There is thus an interesting correlation: phonological properties that are relevant to clitics are also relevant to syntax, those that are not relevant to clitics are also not relevant to syntax. Why? Why is only the more 'syntax-interacting side of phonology' relevant to clitics, and what does it mean that phonology has a more 'syntax-interacting', structural, side?

This empirical picture is very similar to the semantic picture: just like the naive notion of semantics empirically splits into two parts, a more structural part (formal, compositional semantics) deeply

intertwined with syntax, and a more contentive non-structural side which is separate from syntax (lexical, or conceptual, semantics) – so does phonology divide into two parts, a structural side deeply intertwined with syntax, and a contentive (melodic) side, independent of syntax. The cross-modular generalisation about clitics suggests an important strengthening of this formulation: 'deeply intertwined' in fact means 'inside the syntactic module': just like the structural side of semantics turned out to be an integral part of syntax, the structural, rhythmic, side of phonology is in fact part of core syntax. If that is true, the modularity puzzle disappears on the phonological side too: the 'phonological' properties of clitics co-varying with syntax are in fact syntactic properties of clitics.

Since stress is part of the 'structural' side of phonology, this translates into a different way of approaching the relationship between syntactic structure and stress. Just like 'definiteness' or 'tense' have grown to be *bona fide* syntactic entities, the generalisation about clitics suggests that entities such as (word) stress should also be viewed as *bona fide* syntactic objects. Not in the traditional derivative way: traditionally the link between syntax and e.g. stress would be interpretational, derivative, calculated (perhaps in part) on the basis of the syntactic structure. Either directly as in eg. Cardinaletti (1993) and Cinque (1993), or indirectly, as in eg. Selkirk (1984) and many after her.[7] The modularity reasoning about clitics suggests going one big step further: just like definiteness or tense are not 'derived' from syntactic structures but rather have their own position within the syntactic functional projections, rhythmic/structural properties of phonology such as stress are also not interpreted or derived from syntax, but rather they have their home inside syntax, in their own positions in the syntactic fseq, perhaps as a wsP – a word-stressPhrase.

Once seen in this light, it becomes conspicuous that the syntactic fseq already has a projection that is closely related to wsP: FocusP. The relevant phonological properties of clitics come close to lackOfFocusP, i.e. *[-focus]*. If that is so, a clitic is a lexeme whose lexical representation lacks the projection wsP (or has a negative value for FocusP, in the relevant sense).

The suggestion is then that either the nominal FocusP or some projection around it host a feature that is *universally* interpreted as word stress in the same way that DefP is *universally* interpreted with the semantics of definiteness.

**7** These derivative approaches to stress suffer from the same modularity problem: they postulate that stress, or prosody more generally, interfaces with syntax. In a modular world, this means being their own module, and thereby reproducing the paradox that both sides of the interface are universally correlated in the generalisations about clitics, while modularity doesn't allow them to be correlated.

## 5 Conclusion

The fact that some semantic and phonological properties of (syntactic) clitics universally correlate with syntactic behaviour indicates that those parts of semantics and of phonology are in fact syntax. This is a lesson that has been gradually incorporated into syntax during the last three decades. In that vein, the behaviour of clitics suggests that the rhythmic side of phonology has its own dedicated functional projections in very much the same way as many formal semantics properties have acquired their own dedicated functional projections.

## Appendix: Syncretism

The core of the reasoning above has remained largely theory neutral, assuming only strong modularity (its conclusions are phrased in terms of a particular theory, that of functional projections, but the reasoning itself is independent of that.) Once a theory of syncretism is adopted, as is the case implicitly in C&S and explicitly in Nanosyntax (eg. Starke 2002; 2009; Caha, De Clercq, Vanden Wyngaerd 2025), it becomes all but unavoidable to grant word stress its own position in the functional sequence (or more generally, to whatever the phonological properties of syntactic clitics are).

Recall that the same-looking pronoun induces word-stress when used as a strong pronoun, but doesn't when used as a clitic. C&S argued that strong pronouns are clitics with two additional layers of functional projections on top of them. Nanosyntax extends this to a principled theory of syncretism: syncretic readings are one and the same base syntactic structure, with additional layers of functional projections present or missing – those layers expressing the differential semantics of the syncretic readings. If that is so, the fact that a pronoun 'loses' its word stress when used as a clitic derives from the fact that it 'loses' its top functional projection(s) when used as a clitic. Word stress must therefore live in those top functional projections.

More generally, both the apparently semantic and the apparently phonological differences between clitic(, weak) and strong pronouns come from those top layers. Which in turn means that *all* phonological-looking behaviour differentiating between clitic and strong pronouns must live in those top-layers of functional projections inside the strong pronouns. Since word-stress is one such behaviour, it follows that word-stress must be represented inside one of those top layers of nominal structure, included in strong pronouns but absent in clitics. Ergo 'wsP'.

## Bibliography

Caha, P.; De Clercq, K.; Vanden Wyngaerd, G. (2025). *Nanosyntax and the Lexicalization Algorithm*. Oxford: Oxford University Press.

Cardinaletti, A. (1993). "The Affrication of [s] in a Central Italian Variety: Syntactic Conditions on a 'Sandhi' Rule". Belletti, A. (ed.), *Syntactic Theory and the Dialects of Italy*. Turin: Rosenberg & Sellier, 99-129.

Cardinaletti, A.; Starke, M. (1999). "The Typology of Structural Deficiency: A Case Study of the Three Classes of Pronouns". Van Riemsdijk, H. (ed.), *Clitics in the Languages of Europe*. Berlin: Mouton de Gruyter, 145-233.

Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.

Cinque, G. (1993). "A Null Theory of Phrase and Compound Stress". *Linguistic Inquiry*, 24(2), 239-98.

Fodor, J.A. (1983). *The Modularity of Mind. An Essay on Faculty Psychology*. Cambridge, MA: MIT Press.

Selkirk, E. (1984). *Phonology and Syntax. The Relation Between Sound and Structure*. Cambridge, MA: MIT Press.

Starke, M. (unpublished). "The Day Syntax Ate Morphology". EGG Summerschool Seminar (Novi Sad, 22 July - 3 August 2002).

Starke, M. (2009). "Nanosyntax: A Short Primer to a New Approach to Language". *Nordlyd*, 36, 1-6.

Zwicky, A.M.; Pullum, G.K. (1986). "The Principle of Phonology-Free Syntax". *Ohio State University Working Papers in Linguistics*, 32, 63-91.