De-syntacticising Syntax? Concerns on the Architecture of Grammar and the Role of Interface Components*

Aritz Irurtzun
CNRS-IKER
aritz.irurtzun@iker.cnrs.fr

Received: April 3, 2018
Accepted: September 23, 2019

Abstract
This article discusses different ways in which interface components could potentially affect syntax (or what have traditionally been analysed as syntactic phenomena). I will distinguish four types of potential effects that the interface components could have onto syntax: (i) no real interaction, since almost nothing pertains to syntax: everything (beyond Merge) is externalization; (ii) computations at interface components actively affect the syntactic computation; (iii) Properties of interface representations function to inform biases for language acquisition; (iv) interface components impose Bare Output Conditions (legibility conditions) that constrain the range of possible syntactic representations at the interface. I argue that the first two are problematic, whereas the latter two may help us understanding a range of universal and variable phenomena.

Keywords: architecture of grammar; syntax; interfaces; bare output conditions; modularity

Resum. Dessintactitzar la sintaxi? Preocupacions sobre l’arquitectura de la gramàtica i el paper dels components d’interfície

Aquest article tracta diferents maneres en què els components de la interfície poden afectar potencialment la sintaxi (o tradicionalment analitzats com a fenòmens sintàctics). Distingiré quatre tipus d’efectes potencials que els components de la interfície poden tenir sobre la sintaxi: (i) no hi ha interacció real, ja que gairebé res no pertoca a la sintaxi: tot (més enllà de combinar) és externalització; (ii) els càlculs dels components de la interfície afecten activament la computació sintàctica; (iii) les propietats de les representacions d’interfície funcionen per informar els biaixos per a l’adquisició d’idiomes; (iv) els components de la interfície imposen condicions de sortida nua (condicions de llegibilitat) que restringeixen el rang de representacions sintàctiques possibles a la interfície. Jo argumento que els dos primers són problemàtics, mentre que els dos últims poden ajudar-nos a comprendre una gamma universal i variable de fenòmens.

Paraules clau: arquitectura de gramàtica; sintaxi; interfícies; condicions de sortida nua; modularitat

* My deepest thanks to the organizers as well as to the audience of the Workshop on Generative Syntax: Questions, Crossroads, and Challenges held at the UAB. This research was funded by the following projects: PGC2018-096870-B-I00 (MICINN & EAI); FF12017-87140-C4-1-P (MINECO); IT769-13 (Eusko Jaurlaritza); BIM (ANR) and UV2 (ANR–DFG).
Aritz Irurtzun

1. Introduction

This article discusses different ways in which interface components could potentially affect syntax (or what have traditionally been analysed as syntactic phenomena). I will distinguish four types of potential effects that the interface components could have onto syntax:

1. **No real interaction, since almost nothing pertains to syntax: everything (beyond Merge) is externalization (section 2).**
2. **Computations at interface components actively affect the syntactic computation (section 3).**
3. **Properties of interface representations function to inform biases for language acquisition (section 4).**
4. **Interface components impose Bare Output Conditions (legibility conditions) that constrain the range of possible syntactic representations at the interface (section 5).**

The first two conceptions advocate for a de-syntactization of phenomena previously though to pertain to the syntactic component, since they take processes and constructions that were classically thought to pertain to the syntactic component to (i) merely pertain to externalization phenomena, or (ii) be derivative of phonological (or semantic) computations. The latter two, on the contrary, are compatible

---

**Figure 1.** The inverted-Y model of the architecture of grammar (Chomsky 1995).
with the classical (inverted-Y) model of the architecture of grammar (Chomsky 1995) where syntax generates structures that at some point will be spelled out to the linguistic components (PF, LF) that serve as interfaces with the language-external A(rticulatory)-P(erceptual) and C(onceptual)-I(ntentional) systems (Figure 1).

Over the last twenty years a range of works adapted this general architecture of grammar to a more dynamic one where, rather than a single point of SpellOut (transfer), the syntactic derivation unfolds in phases (computational cycles) leading to a spell out of chunks of structure at various points (see i.a. Uriagereka 1999; Chomsky 2000, 2001; Kratzer & Selkirk 2007) (Figure 2).

Even though the syntax of phase-based structure building has been quite extensively explored, there are still very few studies devoted to the nature of phases at the interfaces, and the concepts and primitives employed in each work can be very different and even incompatible with each other (see e.g. Marvin 2003; Dobashi 2003; Kratzer & Selkirk 2007; Samuels 2011; D’Alessandro & Scheer 2015; or the works in Gallego 2012). In any event, there is no substantial architectural difference between the model depicted in Figure 1 and the one in Figure 2 with respect to the derivational relationship between the syntactic computation and the interfaces.

Regarding the four types of potential relationships between syntax and the interfaces, the first conception of the interface constitutes a theoretically attractive program, but I will argue that it requires a large number of specifications and additional theoretical primitives if every single point of cross-linguistic variation is to be conceived as a PF phenomenon (cf. section 2). Furthermore, it also faces non-trivial empirical challenges (namely, the existence of cross-linguistic variation in the available semantic representations).

Then, section 3 argues that the second type of interaction requires a radical change in our conception of the architecture of grammar, as has been extensively explored.

---

1. See also Marušič (2005, 2009) for a model with non-simultaneous transfer to PF and LF.
claimed in the literature (see e.g. Jackendoff 1997, 2002; Zubizarreta 1998; van der Hulst 2006). However, I believe that there is no genuine evidence requiring such a radical change (i.e. that syntax is ‘phonology-free’ (Zwicky & Pullum 1983; Miller et al. 1997; or ‘melody-free’ Scheer 2011). I will discuss what probably seems to be the best case for such an interactive architecture (the focus to stress correspondence, which is argued to underlie focalization movements and wh-constructions Reglero & Ticio 2013, i.a.). I will argue that there is no basis for such a position and that it incurs in a number of conceptual and empirical problems (section 3, see also Irurtzun 2007, 2009). I will also discuss a more recent proposal by Richards (2010) which rather than building on the Nuclear Stress Rule, takes prosodic phrasing to be at the origin of the different interrogative strategies attested cross-linguistically. Based on recent work with M. Duguine (Duguine & Irurtzun 2019), I will raise a number of empirical problems that cast doubt on this vision too.

Now, the vision that interface components reflect to a certain degree the structures generated by the syntactic component seems to be a sensible one; I will discuss frameworks of early language acquisition that illuminate the acquisition of syntactic patterns via prosodic and semantic bootstrapping hypotheses (section 4).

Last, I will discuss how legibility conditions imposed by language-external components (the Articulatory-Perceptual systems at the interface with PF and the Conceptual-Intentional systems at the interface with LF) may affect the design of our syntactic ability. I will argue that alongside other effects, investigation into such ‘legibility conditions’ could help us understanding intriguing linguistic phenomena such as the cross-linguistic lack of verbal wh-words (section 5).

A final section with general conclusions closes the article.

2. Radical externalization

There is a sort of tension in contemporary syntactic theorizing between approaches that seek to explain phenomena at the interface between discourse and syntax as being eminently syntactic (e.g. the cartographic enterprise of Rizzi 1997 and others, or the rich articulation of the discourse-syntax interface in Haegeman & Hill 2014) on the one hand, and more programmatic proposals such as Berwick & Chomsky (2011) or Boeckx (2011, 2014) that argue that syntax is basically just Merge (structure building), and all cross-linguistic variability is restricted to the externalization component, on the other hand. For instance Boeckx (2014) defends the Strong Uniformity Thesis, with the consequence that “all of cross-linguistic variation reduces to realizational options available in the externalization component (‘PF’)” (Boeckx 2014: 139):

(1) **Strong Uniformity Thesis**: Principles of narrow syntax are not subject to parametrization; nor are they affected by lexical parameters.

According to this hypothesis, phenomena that show variable patterns that were previously thought to derive from syntactic parameters are better under-
stood as differences in the realization/externalization of a cross-linguistically homogeneous underlying syntax. This is, in a nutshell, what Tokizaki & Dobashi (2013) and Tokizaki (2016) call the ‘Universal Syntax and Parametric Phonology’ thesis.

This ambitious research program is nonetheless virtually unexplored. The most complete such proposals may be Tokizaki’s (2013) analysis of the compounding parameter as deriving from word-prosodic restrictions, or Mathieu’s (2016) analysis of the variability in the realization of wh-questions (previously analysed under the “wh-parameter”).

What follows discusses these proposals (sections 2.1 and 2.2 respectively) and section 2.3 overviews some of the general empirical problems that radical externalization theories face.

2.1. A radical externalization approach to the compounding parameter

Tokizaki (2013) proposes that the cross-linguistic availability of recursive N+N compounding derives from word-prosodic restrictions:

(2) A complement moves to the specifier position to make a compound if the resulting structure has an acceptable prosody of a word in the language (Tokizaki 2013: 284).

For instance, he argues that the fact that English has productive N+N compounds such as (3) depends on the stress-pattern of the resulting compound (i.e. that the resulting structure has the same stress location as a word in that language (represented in the following examples with the stressed syllable underlined)):

(3) banana-box

In contrast, its Spanish variant in (4) would be ungrammatical because it would have ante-antepenultimate stress ([−4]) in a language that normally has stress on the penultimate [−2] syllable (but that can have stress in any of the ultimate [−1], penultimate [−2] or antepenultimate [−3] syllables):

(4) *banana-caja
   banana-box
   banana-box [Spanish]

The spirit of the idea is notably minimalist, but I think that this analysis may be too powerful, for it predicts that Spanish should allow for compounds, provided that they keep stress in the penultimate position (the ‘default’ stress position in that language). Test cases would be examples such as the ones in (5), which are ungrammatical:
Furthermore, it is not clear how such a proposal could capture the clustering phenomena typically associated to the compounding parameter such as the (un)availability of resultatives (Snyder 1995, 2001). The original observation of Snyder is that transitive resultative constructions such as English (6) are possible only in languages with productive N+N compounding (cf. the typology in Table 1).

(6) John hammered the metal flat.

Thus, I believe that something else should be said about these patterns if we are going to accept that they are due to patterns of externalization.

In the next section I review one of the most detailed and most ambitious proposals for a radical externalization approach to a phenomenon that has generally been thought of as syntactic: Mathieu’s (2016) analysis of the cross-linguistic distribution of different wh-question strategies.

2.2. A radical externalization approach to wh movement vs. wh in situ

Mathieu’s (2016) analysis is concerned with the licensing of wh in situ. It is conceived as a ‘radical externalization’ approach where “the wh parameter is completely relegated to PF” (Mathieu 2016: 252). More precisely, the interrogative strategy/ies that can be used in any particular language depend(s) on its prosodic properties (with respect to the expression of prominence). In this regard, the analysis distinguishes between two types of languages:

(5) a. *sol-luz
sun-light
sunlight

b. *terror-rey
terror-king
king of terror

c. *cristal-cruz
crystal-cross
cross of crystal

d. *champán-bar
champagne-bar
champagne-bar

e. *jazmin-té
jasmine-tea
jasmine-tea

f. *maíz-pan
corn-bread
cornbread
1. **Culminative languages** (e.g. Germanic and (most) Romance): these languages “have lexical stress and always link the prominence of the focused constituent to a stressed syllable” (Mathieu, 2016, 264).

2. **Demarcative languages** (e.g. Korean and Japanese): these languages “resort to the insertion of boundaries either to the left or right (or both) of the intonational phrase to mark focus without any pitch accent on a particular syllable” (Mathieu 2016: 264).

Importantly, such a typological division of languages is taken to be a highly consequential one: Mathieu (2016) argues that *wh*-in-situ languages are languages that use the demarcative strategy only. Thus, “French is a *wh*-in-situ language because of its inherent prosodic properties and in particular because of the way focus is realized in the language. More generally, [Mathieu argues] that, whereas *wh* movement languages tend to use pitch accents followed by deaccenting to express focus, *wh*-in-situ languages tend to use prosodic phrasing. Languages in the first group usually have lexical stress, whereas those in the second one do not. In other words, the option to move or not to move in a given language is constrained by the limits imposed by the phonology of the language. Variation is thus not part of syntax but completely external to it” (Mathieu 2016: 281).

I believe that this approach is mistaken. In recent work (Duguine & Irurtzun 2019) we have argued that such a proposal is problematic; two main types of problems could be mentioned: On the one hand, there is no clear ground for the typological distinction between “culminative” vs. “demarcative” languages. In Mathieu’s (2016) analysis says that “while many languages that use the culminative

### Table 1. The Compounding Parameter and cross-linguistic variation (from Snyder 2001, 329)

<table>
<thead>
<tr>
<th></th>
<th>Resultatives</th>
<th>N+N Compounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Sign Language</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Austroasiatic (Khmer)</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Finno-Ugric (Hungarian)</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Germanic (English, German)</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Japanese</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Korean</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Sino-Tibetan (Mandarin)</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Tai (Thai)</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Basque</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Afroasiatic (E. Arabic, Hebrew)</td>
<td>no</td>
<td>no (?)</td>
</tr>
<tr>
<td>Austronesian (Javanese)</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Bantu (Lingala)</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Romance (French, Spanish)</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Slavic (Russian, Serbo-Croatian)</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>
strategy also make use of the demarcative strategy, the reverse is not true”, but the claim that languages classified as demarcative do not employ pitch accents seems to be unwarranted. In fact, languages classified as demarcative do not restrict their expression of focus to phonological phrasing, but amply employ pitch accents and other local prosodic events to mark focus: higher F0 excursion in pitch accents and tone bearing units, elongated moraic/syllabic duration, higher intensity values, and gestural hyperarticulation are all attested in “demarcative” languages such as Japanese (see e.g. Beckman & Pierrehumbert 1986; Pierrehumbert & Beckman 1988; Fujisaki & Kawai 1988; Maekawa 1999; Kubozono 2007; Venditti et al. 2008; Ishihara 2011, 2015), Korean (Hwang 2006; Lee 2007; Hwang 2011; Kim & Jun 2009), or Mandarin (Xu 1999; Gu et al. 2003; Liu & Xu 2005; Chen & Gussenhoven 2008; Lee et al. 2016). Besides, the existence of the consequential cross-linguistic tendencies with respect to wh-questions is not obvious either: for instance, Amharic has stress-accent (Haile 1987) but also wh-in-situ (Eilam 2008), as do Pashto (Tegey & Robson 1996; David 2014), Uyghur (Yakup & Sereno 2016; Major 2014), Marathi (Wali 2005; Rao et al. 2017; Dhongde & Wali 2009), or Ancash Qechua (Hintz 2006; Cole & Hermon 1994).

Actually, Basque is illuminating in this respect. In general, this language shows syntactic homogeneity across its dialects, but a wide prosodic variability. So it appears problematic for any approach derivationally tying syntactic patterns to phonological patterns; the theory would predict that they should co-vary but often times they do not (see also section (3)). For instance, research into cross-linguistic prosodic typology over the last twenty years has underlined a range of similarities between the word-prosodic patterns of Northern Bizkaian Basque (and just those varieties of Basque) and Tokyo Japanese. Hualde et al. (2002: 578) even argue that “the striking coincidence between some Basque varieties (NB) and Tokyo Japanese in a number of important prosodic properties suggests that this set of common properties can be used to characterize a prosodic prototype: T-type pitch-accent” (see also Elordieta 1998; Ito 2002; Gussenhoven 2004 for discussion). However, Northern Bizkaian Basque is an obligatory wh-movement variety (cf. Hualde et al. 1994), unlike Japanese. On the other hand, one of the few syntactic differences across Basque dialects is to be found on wh-constructions (and focalizations): as a matter of fact, Labourdin Basque is a stress-accent variety (cf. Gaminde & Salaberria 1997; Hualde 1999, 2003), hence a culminating language under Mathieu (2016)’s typology, but it is a variety that has recently developed optional wh in situ (Duguine & Irurtzun 2014), unlike the rest of stress-accent varieties of Basque that have obligatory wh-movement (see Irurtzun 2016, for an overview).

In conclusion, I think that we cannot maintain that syntactic operations such as wh-movement are dependent on phonological computations of stress, since such a theory predicts a co-variation not observed in the cross-linguistic comparison.

2.3. General problems for the radical externalization thesis

Beyond specific proposals, I would like to discuss how the thesis that all crosslinguistic variation is restricted to externalization faces nontrivial problems with
patterns in which syntactic ("word order") variation seems to be correlated with semantic variation. In principle, the prediction of the radical externalization thesis is that we should not observe any cross-linguistic variability in semantics (no matter it is genuinely semantic in essence, or derived from syntax). What follows provides a glimpse into some types of phenomena that are problematic for the radical externalization thesis.

I would like to underline from the outset that even if we discovered that the typological generalizations at the base of the following proposals were not that strong, the existence itself of variation in semantics (the availability or not of a certain reading in some languages/idioms) casts doubts on the viability of a radical externalization thesis. In the following I briefly review three proposals of variation with interesting ties to syntax, but it should be noted that pure idiosyncratic semantic variation itself would also constitute evidence against the radical externalization hypothesis (see for instance the references at the end of this section).

2.3.1. Differences in the interpretation of interrogatives

The first example of nontrivial semantic differences that I would like to discuss concerns the interpretation of multiple wh-interrogatives. Bošković (2003) analyses the patterns of interpretation of wh-movement and wh in situ languages, and his observation is that wh in situ languages allow for both Pair-List and Single-Pair interpretations of multiple wh-question sentences. As an illustration, consider example (7) from Japanese:

(7) Dare-ga nani-o katta no? [Japanese]
    who-nom what-acc bought q
    Who bought what?

Question (7) can be felicitously used in a situation where we have a shopping list for a party and each guest ought to buy something. In such a situation, (7) could be answered with a list of buyers and their corresponding items (say, as in (8)). This is the so-called Pair-List interpretation:

(8) Hanako-ga wain-o katta, Miki-ga biru-o katta… [Japanese]
    Hanako-nom wine-acc bought Miki-nom beer-acc bought…
    Hanako bought wine, Miki bought beer…

A similar thing happens in the English counterpart to (7) in (9A), which can be naturally answered by a list such as (9B):

(9) A. Who bought what?
    B. Mary bought wine, Susan bought beer…

However, Bošković (2003) argues, the difference between wh in situ languages like Japanese and obligatory wh-movement languages like English is that multiple questions like (7) in wh in situ languages also allow for Single-Pair answers,
whereas their counterparts in \(wh\)-movement languages do not. As such, (7) can be felicitously uttered in a context such as the one in (10), which requires a PairList reading:

(10) **Context:** John is in a store and in the distance sees somebody buying a piece of clothing, but does not see who it is and does not see exactly what the person is buying. He goes to the sales clerk and asks the question in (7).

Languages displaying obligatory \(wh\)-movement like English on the other hand, do not have the Single-Pair reading, and as a consequence (9A) cannot be uttered in the context of (10). What is more, optional \(wh\)-movement languages like French provide strong evidence for such a typological claim, with each type of construction patterning as expected. The \(wh\) *in situ* construction of (11a) allows for both Pair-List and Single-Pair readings, whereas the \(wh\)-movement construction of (11b) only allows for the Pair-List reading:

(11) a. *Il a donné quoi à qui?* [French]
   he aux given what to whom
   What did he give to whom?

b. *Qu’a-t-il donné à qui?*
   what.has.he given to whom
   What did he give to whom?

The analysis in Bošković (2003) is that Single-Pair multiple \(wh\)-questions denote sets of propositions (type \(<pt>\)), but pair-list multiple \(wh\)-questions denote sets of questions (\(i.e.,\) sets of sets of propositions; type \(<pt,t>\)), and that it is the overt movement of a \(wh\)-phrase to Spec-CP that generates the loss of the Single-Pair interpretation. Assuming, following Hagstrom (1998), that the interrogative Q morpheme is an existential quantifier over choice functions, its merger outscoping several \(wh\)-phrases will derive in a Single-Pair reading (when the \(wh\)-phrases are left in situ, (12)), whereas movement of a \(wh\)-phrase to C (and crossing Q) generates a relativized minimality effect, which derives in the loss of the \(wh\) *in situ* reading (13):

(12) \(C \ Q [wh_1 \ wh_2 \ V]\)

(13) \(wh_1 \ C \ Q [t \ wh_2 \ V]\)

Questions with Pair-List interpretations can be generated with no problems, for the movement does not affect the scope of the Q particle, which is always attached to the lowest \(wh\)-phrase:

(14) \(wh_1 \ C [t \ wh_2 + Q]\)

See Bošković (2003) for more details on this theory of the syntax-semantics interface but importantly, whatever our analysis of these facts, they seem to provide
strong evidence against the idea that all cross-linguistic variation is restricted to interface components. Here, it seems that there is a categorical semantic variation and that syntax (word order) and semantics are observed to go hand in hand.

In the following, I will briefly mention a couple of similar cases that in my view provide reasons for scepticism with respect to the radical externalization approach.

2.3.2. Differences in the availability of ‘telic pairs’

In a series of works, Higginbotham (2009a, 2009b) discusses the idea that accomplishments like resultative constructions are syntactically represented by ordered pairs of eventualities, and that “the ‘accomplishment’ interpretation of a predicate may stem from the complex thematic structure $<E, E'>$ of a preposition, a syntactic adjunct, rather than from the head verb” (Higginbotham 2009a: 116). His claim is that the structures $<E, E'>$ are telic pairs; holding that the formation of telic pairs is a compositional, rather than a lexical, process. There are languages with the possibility of generating telic pairs like English, where example (15) is ambiguous between a stative and a motion reading (i.e., ‘the boat stays floating under the vertical projection of the bridge’ vs. ‘the boat went to some space under the bridge, floating the while’). However, other languages such as Italian with constructions like (16), only have the stative reading:

(15) The boat is floating under the bridge.

(16) La barca galleggia sotto il ponte. [Italian]

      the boat   under the bridge
The boat is floating under the bridge. (stative)

In order to account for this type of data, Higginbotham (2009a, 2009b) proposed a (de-)compositional analysis whereby the essential difference between languages like English and Italian is that English allows for a combinatorial operation that generates telic pairs of events, whereas Italian doesn’t. Interestingly, the idea in Higginbotham (2009a, 2009b) is that this feature is not idiosyncratic of V-P constructions, and he proposes that the same mechanism underlies complex constructions such as resultatives (17), which are naturally available in languages allowing the motion directional reading like English (or Chinese), but totally absent in languages lacking it such as Italian.

(17) I wiped the table clean.

The proposal in Higginbotham (2009a, 2009b) may raise skepticism, for it proposes a semantic parameter distinguishing languages allowing a specific semantic combinatorial operation and languages disallowing it (see also also Table 1 and Snyder 1995, 2001, 2005 for a proposal on related constructions), but the range of phenomena discussed cannot easily be reduced to a mere externalization parameter: they involve complex syntax-semantics pairings which apparently can be generated in some languages but not in others.
2.3.3. Interpretive consequences of V raising
In recent work Han et al. (2016) have analysed the variability with respect to verb raising observed across Korean idiolects. Korean being a verb-final language, its basic word order (18) is compatible with both verb-raising (19) and tense lowering (20) constructions:

(18) Kim-i cacwu Lee-lul piphanha-n-ta [Korean]
    Kim-nom often Lee-acc criticize-pres-decl
    Kim often criticizes Lee.

Thus, an important part of the input that Korean-learning children are exposed to is critically underspecified as to whether it was generated with a verb raising grammar or a tense-lowering grammar. However, as argued by Han et al. (2007, 2016) the relative scope between negation and object QPs provides an appropriate diagnostic for the position of the verb in a Korean speaker’s I-language: if there is verb raising, negation (a clitic) moves with it, and as a consequence it outscopes the object QP. On the contrary, if there is no verb-raising, the object QP takes scope over negation.

The preceding literature on the topic provided mixed judgments on these issues and a blurred theoretical image, but Han et al. (2016) show that rather than a stochastic procedure, the option of V raising vs. T lowering is grammaticalized in each Korean idiolect, and that there are actually two varieties of Korean grammar coexisting: one with verb-raising (19), the other one without it (20). Remarkably, the participants in Han et al.’s (2016) study show stable judgments across test items and experimental sessions.2

2. Furthermore, the rarity and scarcity of the evidence that signals whether a chain was generated with a verb-raising or a tense-lowering grammar makes that this grammatical choice is left underspecified so that children do not necessarily converge on the same grammatical options of their parents. As a consequence, children’s grammar is not necessarily the same as their parents’, as illustrated by the lack of correlation between their judgments. Han et al. (2016) suggest that this is a case of ‘endogenous’ linguistic variation.
Again, concerning our discussion in this paper, I take it that the fact that this virtually invisible movement has predictable and stable semantic consequences argues against the conception that all variation is restricted to the externalization component.

2.4. Conclusion

The radical externalization thesis is an elegant programmatic position that seeks to understand the commonality to human languages (actually, to human language) on the basis of the idea that syntax (structure building computation) is homogeneous across the species and that it is inherently directed towards the construction of complex thought. Externalization would just be secondary to this process, and all the cross-linguistic variability would arise during externalization (it would amount to different ways of expressing the same structure/thought).

I believe that the evidence that I have discussed casts doubts on such a position. I would like to underline again that one does not have to believe in the reality and generality of proposed ‘semantic parameters’ such as Chierchia’s (1998) parameter for the differential denotation of nominals across languages, or the aforementioned ‘semantic composition’ parameter by Higginbotham (2009a); see for instance Duguine et al. (2017) for discussion. If there are non-trivial semantic differences across languages, given the inverted-Y model of the architecture of grammar they can only derive either from syntax (whereby different positions determine differences in interpretation (say, different landing sites determine different scopal interpretations)) or from semantics itself (the use of different semantic combination operations, different domains, etc.). They cannot derive from externalization if externalization has nothing to do with the path to LF/SEM.

The discussion above briefly commented on a couple of cases that cannot be captured in terms of externalization in any obvious way, but the literature on the syntax-semantics interface is full of similarly problematic phenomena of cross-linguistic variation (see i.a. Bach et al. 1995; Chung & Ladusaw 2004; Matthewson 2010; Arregui et al. 2014; Matthewson 2014; Etxeberria & Giannakidou 2014; Holmberg 2016; Keenan & Paperno 2017; Scontras et al. 2017). It is unlikely that this type of phenomena can be accounted for as differences in externalization.

Next section discusses another type of conceivable relationship between interface components and syntax; namely, the hypothesis that interface components may interact with syntax during derivations.

3. Actively affecting the syntactic computation

A more active way in which interface components may affect ‘syntactic’ computations is by having parallel computations in, say, phonology and syntax where structures generated in the former serve as the structural description for the operations taking place on the latter.
However, this type of proposals require a radical change in our conception of the architecture of grammar, for in the ‘inverted-Y’ model interface components cannot directly interact with syntax (Zwicky & Pullum 1983; Miller et al. 1997; Irurtzun 2007, 2009). Proponents of this type of interactions have thus proposed alternative conceptions of the architecture of grammar allowing such types of interaction. For instance, the one in Figure 3, from Vallduví (1995), presents a direct link between SS and LF, and an indirect link with PF, with an additional submodule of Information Structure (IS) which is somehow parallel to the syntactic computation (the dashed line means that further strata may be needed to represent other relations) (Figure 3).

More famously, Zubizarreta (1998) proposes a different type of architecture with the level of LF at the center stage of the derivation (Figure 4).

As can be seen, this conception is quite in line with the ‘radical externalization’ hypothesis presented in the previous section. According to Zubizarreta’s (1998) model, the derivation unfolds creating sets of phrase markers until one single phrase

![Figure 3. A model of grammar that incorporates a separate level of IS (from Vallduví 1995: 147).](image3)

![Figure 4. A model of grammar with a post-LF level of Assertion Structure (from Zubizarreta 1998: 32).](image4)
marker is obtained at the level of $\Sigma$-Structure. At this point, operations such as Focus Marking, the Nuclear Stress Rule (NSR) and prosodic movements take place until we reach the level of LF. There the derivation branches in two branches, one that derives in a PF representation and the other one in the “Assertion Structure”, which is the information structure of the sentence where the focus-presupposition partition is encoded.

Last, works like Jackendoff (1997, 2002) have proposed an even more powerful model with fully parallel phonological, syntactic and semantic components with independent generative power that generate structures that are then linked via structure interface (or correspondence) rules (Figure 5).

Probably the best candidate for phonology affecting syntax may be the purported correspondence between focus and stress, which is taken to drive movement operations with semantic consequences in some languages. In the next section, I briefly overview the major tenets and some shortcomings of such approaches. Next, in section 3.2 I analyse a novel take on $wh$-questions that is based on the same type of conception of the architecture of grammar, since it builds on p-phrasing for explaining the $wh$-question construals available crosslinguistically. I argue that this type of proposals have important shortcomings.

3.1. The Nuclear Stress Rule and focus/$wh$-questions

The Nuclear Stress Rule (NSR) governs the assignment of nuclear stress in the clause. The classical theory of Halle & Vergnaud (1987) was a variable stress assignment rule with different parameters such as “head terminal” (“whether or not the head of the constituent is adjacent to one of the constituent boundaries” (Halle & Vergnaud 1987: 9), and “BND” for “boundedness” (“whether or not the head of the constituent is separated from its constituent boundaries by no more than one
“intervening element” (Halle & Vergnaud 1987: 10). For English, the NSR would have the parameter setting in (21):

(21) **The Nuclear Stress Rule** (Halle & Vergnaud 1987: 264):
— The Parameter settings on line N (N≥3) of the Metrical Grid are [-BND, +HT, right].
— Interpret boundaries of syntactic constituents composed of two or more stressed words as metrical boundaries.
— Locate the heads of line N constituents on line N+1.

With this setting, the nuclear stress assignment to Judea in (22) is explained as a simple bottom-up composition of the metrical grid (ex. 83, p. 265 of Halle & Vergnaud 1987):

(22) Jesus preached to the people of Judea.

```
Line 6: (Jesus [preached to the [people of Judea]])
Line 5: (Jesus [preached to the people of Judea])
Line 4: (Jesus [preached to the people])
Line 3: (Jesus [preached])
```

However, in one of the most influential articles on the syntax-phonology interface, Cinque (1993) argued that the phonological parametrization of the NSR was superfluous, for it missed a generalization: nuclear stress is crosslinguistically assigned to the most deeply embedded element with the syntactic structure, so at the interface it suffices to turn syntactic phrase structure into phonological metrical grids and the more embedded an element is in the syntax, the more embedded it will get in the phonology:

(23)

As a consequence, the positional variability observed in head first (SVÓ) vs. head last languages (SÓV) is illusory: it is not the case that in SVO languages such as English or Spanish nuclear stress is assigned to the last element vs. in SOV languages such as Japanese or Basque it is assigned to the central element. In fact, in both types of languages it is assigned to the most deeply embedded element (which
happens to be the O in the unmarked case). Then marked operations such as stress shift or prosodically motivated movements will take place in order to guarantee that the element to be focused receives nuclear stress in PF.

Built upon these observations, a whole line of analysis developed seeking to account for the patterns observed in focus (see in particular Zubizarreta 1998; Reinhart 2006) and interrogative (Reglero & Ticio 2013) constructions as deriving from a purported PF constraint requiring the element to be interpreted as focus to get nuclear stress. Thus, the movements observed in these constructions in some languages are taken to take place in order to guarantee that the focus/interrogative is placed in the most embedded position (the position where it will get nuclear stress by the NSR).

I think that this type of approach is misleading: in previous work (Irurtzun 2007, 2009). I have argued that such a position is not tenable given that it faces a range of conceptual and empirical problems. Here I will not repeat those arguments but I would like to stress a couple of points, based on recent discussions on the literature.

One of the key assumptions of the NSR-based theory of focus is that nuclear stress is not just a correlate of focus; nuclear stress is taken to be not just one of the possibly many manifestations of an underlying focus representation, it is rather—according to this theory—an essential part of the nature of focus, so much so that the whole derivation is affected so that nuclear stress ends up being assigned to a specific item. This alleged intimate relationship between nuclear stress and focus could be understood in an embodied cognition approach as a grammaticalization of the ‘effort code’ (Gussenhoven 2004): more articulatory effort amounts to more vibration of the vocal folds (articulatory phonetics), which in turn amounts to higher excursion in f0 frequencies (acoustic phonetics), which in turn corresponds to a categorical distinction in terms of pitch accent (phonology), and which finally is associated to a contrastive or emphatic interpretation (semantics), that is, focus.

However, there are many languages that do not behave as suggested by this vision. One such case is Mandarin, where items lexically associated to Tone 3 (a falling tone) show lower values when pronounced under focalization (Lee et al. 2016). That is, even if Tones 1, 2, and 4, which involve f0 rises, display higher f0 values when pronounced in focus, Tone 3 does the opposite and reaches lower f0 values when contrastively focused. This casts doubt on the assumption that PF demands focal elements to be associated to higher f0 values. A potential way to circumvent the problem posed by this type of evidence would be to say that the Mandarin data could be taken to indicate that the PF demand is really to somehow ‘hyperarticulate’ the focal element, so that when its tone involves phonological rises, higher f0 values are obtained, and when it involves phonological falls, lower f0 values are obtained. In a nutshell, it would be a matter of exaggerating the tonal events so that the signal involves a larger f0 excursion and the overall acoustic pattern is clearer/easier to discriminate.

Intuitive as this line of thought may be, I think it is not correct. To begin with, other languages such as Akan (Kügler & Genzel 2011) employ pitch register lowering to signal focus, which argues directly against the purported condition on hyperarticulation. In Akan, as in Mandarin, L tones are pronounced with lower F0 values
when in focus, but the same strategy is employed with H tones too; and the more emphatic the interpretation of the focus, the lower the pitch register both for L tones and H tones. Last, other languages do not employ any prosodic means for marking focus. Such is, for instance, Malay, as reported by R. Maskikit-Essed and C. Gussenhoven on a paper illustratively entitled “No stress, no pitch accent, no prosodic focus: the case of Ambonese Malay” (Maskikit-Essed & Gussenhoven 2016), but see also among others Zerbian (2007) on Sotho, Downing (2007) on Chichewa, Chitumbuka and Durban Zulu, Kügler & Skopeteas (2007) and Gussenhoven & Teeuw (2008) on Yucatec Maya, Gut et al. (2013) on Malaysian English, Wang et al. (2011) on Wa, Daeng (Mon-Khmer) and Yi (Sino-Tibetan), or Xu et al. (2012) on Taiwanese Mandarin.3

3.2. P-Phrasing and the interrogative strategies

In recent work, Richards (2010) (see also Richards (2016)) proposes a theory according to which the interrogative strategies used by specific languages are (in part) determined by their prosodic properties, but instead of basing his analysis on nuclear stress placement, he builds it on the idea that at PF there is a constraint requiring the wh-word and the interrogative complementizer to be contained within the same prosodic phrase:

(24) “Given a wh-phrase α and a complementizer C where α takes scope, α and C must be separated by as few Minor Phrase boundaries as possible, for some level of Minor Phrasing” (Richards 2010: 151).

Regarding cross-linguistic variability, different languages are said to satisfy this constraint by appealing to different strategies:

— Changes in the prosodic phrasing (some sort of “Prosodic rephrasing”).
— ‘wh-movement’ to the C domain.

In line with minimalist desiderata, the idea is to derive the question formation strategy that a language will employ from parametric choices which are independent of question-formation. These would be (i) the relative order of heads and their complements (locus of C°), and, crucially, (ii) the alignment pattern of prosodic phrase boundaries.

Within this system, Richards (2010) analyses (Northern Bizkaian) Basque as a variety with final complementizers and Minor Phrase Boundaries to the right of certain XPs such as wh-phrases, studying the patterns of wh-questions in Basque as

3. Beyond the range of problems discussed in Etxepare & Uribe-Etxebarria (2005, 2012), the analysis of Reglero & Ticio (2013) linking wh-phrases to nuclear stress faces a further problem in that the purported cross-linguistic association between wh-words and nuclear stress is not cross-linguistically stable (the case of Italian is famous for instance (Ladd 1996; Bocci et al. 2017), cf. also the observation that during language acquisition children may tend to directly drop wh-phrases (De Lisser et al. 2015).
being derivative of a dynamic syntax-phonology interface that seeks to guarantee the fulfilment of the condition in (24).

It should be noted that approaches tying ‘syntactic’ structures and phonological structures can and do face granularity problems; in principle, they predict that prosodic typologies and syntactic typologies should co-vary. Restricting my discussion to Basque, Richards’s (2010) analysis cannot be extended to other dialects of the language which, having a different word-prosodic system, have the very same syntax for interrogatives (say, any of the stress-accent varieties, from Central Basque to Souletin (Irurtzun 2016)). It cannot either account for the emergent Labourdin variety which, having a similar prosodic system to those of other varieties, has a different syntax for interrogatives.

Once again, I completely agree with the desirability of seeking extra-syntactic explanations for syntactic patterns, but I am afraid that those theories are too powerful and too weak at the same time. They predict co-variation between syntax and phonology that is not observed cross-linguistically, and the models of architecture implied in those works require substantive changes if they are to obtain explanatory power.4

In the next two sections I discuss two further possible ways for the interfaces to affect syntax that, I’ll suggest, may indeed help us understanding patterns of syntactic universals and variation.

4. Reflecting syntax and biasing acquisition

Contrasting with the previous conceptions, a different line (or lines) of investigation hypothesize that interface components have an important impact not in syntactic derivations but on the development of the syntactic hypothesis space that a child will consider during early language acquisition. The idea is that the child uses perceptual input (visual/acoustic signals and situations) to hypothesize grammatical structures during language acquisition, assuming some degree of homomorphism between syntactic structures and the representations of the input at interface components.5 Thus, a relative homomorphism between syntax and the interfaces, combined with the sensorial experiences of the children serve to bias the process of language acquisition.

Within this general sets of ideas, two main areas of research have been developed:

— Semantic bootstrapping theories for predicate argument structure.
— Prosodic bootstrapping theories for head-complement orders.

In the following, I briefly present the major tenets of each of these approaches. The argument will be that these approaches uncover processes where semantic and

4. See Duguine & Irurtzun (2019) for further criticisms.
5. One may conjecture that the relative homomorphism derives from economy/simplicity metrics in inter-modular interface transductions (cf. Reiss 2007; Graf 2013).
phonological information serve to inform the determination of what traditionally have been taken to be syntactic phenomena (patterns of phrase structure and word order, and argument structure). Note, however, that this is radically different from having interface components actively affecting syntactic derivations.

4.1. On the LF side: semantic bootstrapping

A major early contribution of the generative enterprise is the idea of the autonomy of syntax (Chomsky 1955: 1957). This hypothesis has an important implication regarding language acquisition, as has been emphasized by Grimshaw (1981) among others: given that there is no deterministic co-variation between syntactic types and semantic types, a child cannot directly deduce a syntactic analysis from an analysis of the semantics of a phrase. As a consequence of this, she must learn the two kinds of information separately. But contrary to what may appear at first sight, this has as a consequence the virtue of easing acquisition. In Grimshaw’s (1981: 169) terms “if there are n bits of syntactic information to be acquired, and m bits of semantic information, n + m bits of evidence are needed for learning in the autonomous theory, nm in the nonautonomous theory”.

The semantic bootstrapping hypothesis can be defined as built on the idea that “the child can access a structural representation of the intended semantics or conceptual content of the utterance, and that such representations are sufficiently homomorphic to the syntax of the adult language for a mapping from sentences to meanings to be determined” (Abend et al. 2017: 117). For instance, “if children know that a word refers to a thing, they can infer that it is a noun; if they know that X is a predicate and Y is its argument, they can infer that X is the head of a phrase that includes Y; if they know that a phrase is playing the role of agent, they can infer that it is the subject of the clause” (Pinker 1989: 425).6,7 Thus, authors like Pinker (1989) have proposed models of language acquisition where the ‘linking problem’ is partially solved via a range of semantically informed hypotheses about the syntax of the elements in the input.

If this hypothesis is correct, semantic information would have an effect on syntax, not directly in the derivational computation, but in biasing development (Figure 6).

4.2. On the PF side: prosodic bootstrapping

The inverted-Y architecture of grammar (Figure 1) makes the claim that syntactic representations are somehow mapped onto prosodic representations. An oneric image of the syntax-phonology interface would give the image of a perfect mapping, such as the one in (25):

---

6. See also Grimshaw (1981); Bates & MacWhinney (1989); Clahsen et al. (1994) and Gleitman (1990) for discussion.
7. See also Markman (1992) for related issues on the problem of induction in word learning for objects.
Figure 6. An idealization of the start of language acquisition according to the semantic bootstrapping hypothesis (from Pinker 1989: 426).

(25)
But even if the empirical reality differs substantially from such a picture, this is the starting point of virtually all analyses of the syntax-phonology interface: the assumption is that there is an interface procedure so that syntactic representations are mapped to (wrapped in, aligned/matched with, etc.) prosodic units (see Nespor & Vogel 1986; Selkirk 1986; Truckenbrodt 1995, 1999; Seidl 2001; Dobashi 2003; Wagner 2005; Tokizaki 2008; Elordieta 2008; Selkirk 2011; Selkirk & Lee 2015, for discussion and a range of different views). In the unmarked (most faithful) case, it will be a direct mapping (XP→φ), but very often purely phonological constraints concerning p-phrase uniformity, symmetry, or minimum and maximum size also come into play, and the result of the interface transduction deviates from the perfectly homomorphic pairing.

With respect to our main discussion here, a number of works have identified interesting patterns of correspondence between PF and syntax with respect to rhythm and word order. In particular, several authors propose that the rhythmic pattern of a language is not an idiosyncratic and isolated property, but rather that it is strongly correlated with word order (i.e., that there are correlations between rhythmic patterns and syntactic patterns in that languages tend to cluster with the same rhythmic and syntactic properties, conforming cross-modular linguistic typologies). Furthermore, the explanation of this typological clustering is proposed to derive from the fact that rhythmic patterns serve to bootstrap the acquisition of the specific syntactic patterns of each language (cf. i.a. Mehler et al. 1988; Christophe et al. 2003; Bernard & Gervain 2012; Gervain & Werker 2013; Langus & Nespor 2013).8

In a nutshell, the basic idea of the prosodic bootstrapping hypothesis is that the relative order between heads and their complements is strongly correlated with the rhythmic type of the language, and that infants use their accumulated knowledge about the prosody of their target language(s) to build informed guesses about their corresponding syntactic pattern. This theory builds on a number of experiments that have shown that languages whose correlates of phrasal accent are increases in duration and intensity tend to be head-initial (with a Verb-Object word order) whereas languages that realize stress through a combination of higher pitch and intensity (and possibly also duration) tend to be head-final (with an Object-Verb word order). This generalization is known as the ‘iambic-trochaic law’ (cf. i.a. Hayes 1995; Nespor et al. 2008; Shukla & Nespor 2010), and is taken to be a basic law of grouping based on general auditory perception (i.e. not specific to language). This law states that units (language or music) that differ in intensity tend to be grouped as constituents in which the most prominent element comes first, whereas units that differ in duration are grouped as constituents in which the most prominent element comes last. As Nespor et al. (2008) put it, “if [their] proposal is on the right track, one of the basic properties of syntax can be learned through a general mechanism of perception”. Summarizing then, the prosodic

8. See also Donegan & Stampe (1983, 2004) who on independent grounds propose a ‘holistic typology’ based on rhythmic grounds in order to account for the polarized structural divergence of languages such as Munda and Mon Khmer.
bootstrapping hypothesis claims that beyond the observed typological correlation between prosodic and syntactic patterns, there is a causal developmental connection between them: babies use prosody to inform their guesses about the syntactic pattern of their target language.9 In favor of this hypothesis, recent studies such as Gordon et al. (2015) suggest that there is a correlation between rhythm perception skills and morphosyntactic production in children with typical language development, others such as Flaugnacco et al. (2014); Leong & Goswami (2014) also argue for a strong association between reading skills and meter perception and rhythm processing, and yet other studies such as Zumbansen et al. (2014) report the beneficial effects of both pitch and rhythm in the clinical therapy for patients with Broca’s aphasia. So, in a nutshell, if the prosodic bootstrapping hypothesis is correct, it would be a case where PF influences syntax, not in derivational terms, but in developmental ones.

5. Legibility conditions at the interfaces

Last, I would like to discuss another possible relation of the interface components with respect to syntax; that of bare output conditions as restrictions on the types of representations that they can handle (and as a consequence, of the types of representations that syntax can provide as its output). In what follows, I will briefly discuss a couple of restrictions that interface components may impose on the output of syntax. They are based on minimal requirements that derive from the architecture of the language-external systems that interface with the linguistic levels of PF (the Articulatory-Perceptual apparatus) and LF (the Conceptual-Intentional apparatus), cf. Figure 1. The general idea is that legibility conditions imposed by language-external apparatus constrain the types of representations that may derivationally arrive there, and that this is reflected in the restricted cross-linguistic variability.

5.1. On the PF side

The nature of the human Articulatory-Perceptual apparatus dictates a range of legibility constraints on the representations it can handle. Arguably, one such case could be the existence of maximum size constraints in prosodic phonology (see i.a. Delais-Roussarie 1995; Selkirk 2000, 2011; Elordieta et al. 2005; Jun 2005), with the result that prosodic phrases tend to be contained within the limits of breath groups (that is, even if in principle computable by UG, phonological p-phrases larger than n syllables would be difficult to produce, and difficult to process and acquire too).

9. Developmentally, such a theory is reinforced by the fact that a large part of the neurocognitive machinery required for processing and learning prosodic patterns is developed before the syntactic abilities mature (potentially, after the postnatal development of a globular brain Boeckx & Benítez-Burraco 2014; Irurtzun 2015).
A syntactically\textsuperscript{10} more interesting case could be the linearity requirement at the Articulatory-Perceptual interface, which would conceivably derive from the nature of human articulators which require to externalize terminal elements sequentially. As a consequence, syntactic trees (which are characterized by phrase structural relations such as dominance, sisterhood, c-command, etc.), have to be linearized for externalization. Here Kayne’s (1994) Linear Correspondency Axiom (LCA) is a well-known procedure for linearizing structures: asymmetric c-command is mapped into linear precedence (but see the work of Biberauer \textit{et al.} (2014) referred to in the previous section). Thus, the tree in (26) is mapped onto the string \textless j\textasciimacron{}m\textasciimacron{}p\textgreater : 

\begin{equation}
(26)
\end{equation}

The LCA has the following three properties (i) it is transitive (if \(xLy \& yLz \rightarrow xLz\)), (ii) it is total (for all \(x,y\), either \(xLz\), or \(yLx\)), and (iii) it is antisymmetric: \(\neg(xLy \& yLx)\). Therefore, the relative precedence orders are \(<J,M>, <J,N>, <J,P>, <M,P>\), for non-terminal elements, and \(<j,m>, <j,p>, <m,p>\) for terminal elements. Given the properties I just mentioned, this is mapped into a linearization of \textless j\textasciimacron{}m\textasciimacron{}p\textgreater .

But interestingly, the last property (the antisymmetric requirement) has important consequences for what has traditionally been analysed as syntactic displacements (\textit{cf.} \textit{i.a.} Chomsky 2016). Movement (internal merge) is taken to generate a copy of an element in a higher position in the tree, creating a new c-command relationship that will generate a conflicting representation at the A-P interface, as represented in (27):

\begin{equation}
(27)
\end{equation}

\textsuperscript{10.} Assuming that the ‘head parameter’ reflects some underlying syntactic difference across-languages.
The tree structure in (27) has two copies of the element XP: the one within YP is c-commanded by Z and the highest copy of XP, which in turn c-commands Z. Thus, at linearization such a representation will derive into conflicting word order requirements (since XP should precede and follow both Z and XP itself): <XP,Z,XP>. The solution natural language has for treating such paradoxes is chain reduction, the deletion at PF of all but one copy (in general, the highest one) such that the structure can be properly linearized without ordering conflicts (see Nunes’s 2004 elegant work on this). However, the important thing here is that the satisfaction of this formal requirement is anti-functional regarding communication: it generates filler-gap dependencies. And this seems to be the general case; as Chomsky (2016: 22) puts it, “[t]he interesting cases are those in which there is a direct conflict between computational and communicative efficiency. In every known case, the former prevails; ease of communication is sacrificed”.

5.2. On the LF side

We can build informed conjectures about the constraints and expectable legibility conditions of the AP interface (insofar they derive from physical properties of our articulators), but the interface between language and the CI interface is much more obscure. In fact, we know much less about the general architecture of our cognition and the properties we may expect it to demand to its inputs, so any proposal with regard to this area is highly speculative. Nonetheless, I believe that by exploring this area too we can advance in the understanding of a range of puzzling phenomena.

For instance Hurford (2007) proposes that the fact that natural language predicates are restricted to taking (at most) four arguments may be a reflex of human constraints for the representation of a single thought, which in turn derives from our ancient visual-attentional system, which only allows to keep track of a very limited set of objects in a given scene and gives rise to our limits for ‘subitization’ (the capacity for recognizing at a glance how many objects are in a group, without verbal counting (Kaufman et al., 1949)) or for visual object tracking (Pylyshyn 2000) among others.

Another possible case for which we could hypothesize an extra-linguistic origin could be a restriction on vacuous quantification (Chomsky 1982; Kratzer 1995), which rather than an essentially syntactic constraint (Potts 2002) could be conceived as deriving from the logical properties of our language of thought. If anything like this is on the right track, we could say that this restriction has as a reflex in the type rigidity imposed on quantifier expressions so that if they fail to bind a variable, the sentence is ungrammatical.

The last example that I would like to discuss concerns a puzzling typological gap. In Irurtzun (2019) I have explored the possibility that a consequential constraint on the logic of predication may help us better understand the prima facie puzzling gap of the lack of genuine verbal interrogative words. The observation is that cross-linguistically, we can ask questions about different participants in the event (subjects, direct objects, or indirect objects (28)), or modifiers of different kind (29):
(28) a. *Who kissed Mary?*
   b. *Whom did John kiss?*
   c. *Who did John give a kiss to?*

(29) a. *Where did John kiss Mary?*
   b. *When did John kiss Mary?*
   c. *How did John kiss Mary?*
   d. *Why did John kiss Mary?*

However, we cannot directly ask questions on the nature of the eventuality itself. That is, there is simply no interrogative pro-verb, so that we can ask questions such as (30):

(30) *Whxyzed John Mary?*

‘What type of event happened such that it has John as external argument and Mary as internal argument?’

The ban on interrogative pro-verbs has seldom been discussed in linguistics. Hagège (2008) only classifies 28 languages as having the property of displaying interrogative pro-verbs (see also Idiatov & van der Auwera 2004), but many of them are not pro-verbs questioning eventuality types, and if they are, they are syntactic and semantically very restricted (see Irurtzun 2019 for discussion). My argument in that work is that the lack of verbal *wh*-words derives from a legibility constraint at the interface between the linguistic computation and the language-external Conceptual-Intentional systems. I depart from the assumption that at LF sentences are Neo-Davidsonian descriptions of eventualities (*cf.* i.a. Parsons 1990; Hornstein 2002; Pietroski 2005) whereby example (31a) gets the logical form representation in (31b):

(31) a. Brutus stabbed Cæsar.

   b. ∃e [Agent(e, Brutus) & Stabbing(e) & Patient(e, Cæsar)]

My argument is that the lack of verbal *wh*-words derives from a general constraint on the logic of predication: predication is characterized by a logical assertoric force whereby a property is ascribed/attributed/applied to an object (*cf.* i.a. McGinn 2000; Burge 2007; Liebesman 2015) and this is incompatible with querying that very same property (just like asserting and questioning are different speech acts). In other words, predicates predicate and it is therefore that predication qua interrogation is incongruent: the logical act of predication cannot be identical to the logical act of querying and as a consequence, natural language allows for questions such as (32a) or (32b), but not for questions such as (32c):
(32) a. $\exists e [\text{Agent}(e,?) \& \text{Stabbing}(e) \& \text{Patient}(e, \text{Cæsar})]$  
   ‘Who stabbed Cæsar?’

b. $\exists e [\text{Theme}(e, \text{Cæsar}) \& \text{Dying}(e) \& \text{Location}(e, ?)]$  
   ‘Where did Cæsar die?’

c. $*\exists e [\text{Agent}(e, \text{Brutus}) \& ?(e) \& \text{Patient}(e, \text{Cæsar})]$  

Besides, an LF along the lines in (32c) would still be unwarranted, since an interrogative predicate like ?(e) crucially devoids the eventuality of any nature (it is completely undetermined), and as a consequence the DPs get no $\theta$-role (as represented in (33)), given that $\theta$-roles directly depend on the structure of the eventuality (cf. Pietroski 2005; Borer 2005; Ramchand 2008). And failure to assign $\theta$-roles violates the $\theta$-criterion (i.a. Chomsky 1981):

(33) $*\exists e [_____ (e, \text{Brutus}) \& ?(e) \& \text{Past}(e) \& _____ (e, \text{Cæsar})]$  

As can be seen, the logical form in (33) is critically underdetermined where (e,Brutus/Cæsar) may correspond to any theta role (agent, experiencer, possessor…).

In a nutshell then, based on the minimal assumption that logically predication is incompatible with interrogation, the lack of verbal question-words that stand for any eventuality type derives directly from the LF illegibility they would generate: their semantics would require predicates predicating and interrogating at the same time and a failure to assign $\theta$-roles to eventuality participants (which, by hypothesis, corresponds to an illegible representation for the CI interface).11

6. General conclusions

This article sought to discuss possible ways of interaction between the syntactic component and the interfaces. These are essential issues that are seldom explored in and of themselves, even if they have been at the core of theoretical discussions over the last half century.

The complexity of many relevant phenomena implies that simplification in one area requires complexity in the other if we are to obtain even descriptive adequacy. In this regard, the latest ‘radical externalization’ thesis seems to me an interesting thesis worth exploring, but I am afraid that for a number of the relevant phenomena, a radical externalization implementation would require an exponential complexity in the externalization path (involving a range of operations that do not look externally motivated). Furthermore, I believe that one of the most pressing problems that this thesis faces is the fact that cross-linguistic variability does not

11. The analysis in Irurtzun (2019) makes a further prediction: the impossibility should be extendable to other analogous elements whose semantic contribution is the introduction of a predicate of events. In fact, this seems to be the case, as shown by the apparent cross-linguistic lack of interrogative adpositions or tense markers.
seem to be restricted to PF; a variety of cross-linguistic semantic differences has been attested in the literature, which casts doubts on the main premise of the radical externalization thesis.

On the other hand, the more interactive approach that makes interface computations actively affect syntactic operations is too weak and too powerful at the same time. It is too weak in that it is unable to account for many of the phenomena that we observe at the interfaces (for instance, it pretends to generate metrical grids or p-phrases before or independently of syntax, but these approaches are never explicit as to how to do so). But it is also too powerful, for it predicts the possibility that syntactic computations may depend on phonological processes, which is unattested cross-linguistically. This is a fact that in my view strengthens the validity of the restrictive inverted Y-model of grammar.

However, such an architecture does indeed allow for the interface components to affect syntax in some ways: on the one hand, during early acquisition infants may base a range of syntactic hypotheses based on already established phonological and semantic knowledge (as has been previously argued for by the prosodic and semantic bootstrapping hypotheses). On the other hand, legibility conditions imposed by the extra-linguistic systems that language interfaces with restrict the range of possible outputs of the syntactic computation. This type of analysis may help us better understanding grammatical patterns without necessarily conceiving ad hoc constraints to those effects. I believe that investigation into these conditions is a promising venue of research that deserves further explorations.

References


[<https://doi.org/10.1162/ling_a_00195>]


[<https://doi.org/10.1080/10489223.2015.1115049>]


[<https://doi.org/10.1515/9783110179897>]


[<https://doi.org/10.1353/lan.2014.0006>]


https://doi.org/10.3389/fpsyg.2016.01058

https://doi.org/10.3389/fnhum.2014.00096

https://doi.org/10.1093/mind/fzu182

https://doi.org/10.1159/000090090


https://doi.org/10.1017/s0952675716000154


https://doi.org/10.3765/sp.3.9


https://doi.org/10.1016/0010-0277(88)90035-2
<https://doi.org/10.1017/S002226796006305>


<https://doi.org/10.1016/S1364-6613(00)01477-7>


<https://doi.org/10.1515/tlr-2013-0018>


Zumbansen, Anna, Isabelle Peretz & Sylvie Hébert. 2014. The combination of rhythm and pitch can account for the beneficial effect of melodic intonation therapy on connected speech improvements in Broca’saphasia. Frontiers in Human Neuroscience 8: 592.