

# Accommodating Phonological Structure: Extrapolation, The EPP and Trace Effects

Luke Madison Smith

January 19, 2016

## Abstract

While Minimalist Theory and the Bilingual Program have striven to understand the syntax of language as an economical and ‘optimally designed’ system, theoretical syntax has, over the past decades, accumulated a wide range of descriptive principles, exceptions and ‘glitches’ that make the content of ‘Universal Grammar’ seem arbitrary and over-encumbering.

Here I argue that many of these principles are emergent from the *phonological* structure of language and should be modeled as such. I argue that irregular phrasal ordering is prosodic in nature, and constraints on the presence of empty categories (that-trace effects the EPP) are more plausibly thought of as phonological ones. Lastly, I take issue with the classical idea of autonomous syntax.

## 1 Introduction

### 1.1 A Forestry Problem

It’s generally appreciated that there is a general theoretical kinship between syntax and phonology. Formally speaking, both are modeled as having analogical primitives (van der Hulst, 2005) and importantly, hierarchical structure (visually represented as the trees we all know and love).

The important caveat is, however, that there *is* a fundamental disconnect between syntactic and phonological trees. In the abstract, a syntactic tree is infinitely recursive: there is no limit on the ‘depth’ of a syntactic derivation. There may be  $n$  many hierarchical levels to a clause, with an infinity of dependencies and subordination (Figure 1).

The phonological system, however, is not necessarily infinitely deep. Utterances are conventionally divided into intonational phrases (IP), then divided into prosodic phrases ( $\varphi$ ), then prosodic words ( $\omega$ ), then feet (F) (Figure 2). Some languages may plausibly warrant another level in between, but it seems to be the case that the human voice can only modulate in such a way as to characterize a finite, in fact, very small set of phonological levels. This is explicitly formulated in the Strict Layer Hypothesis (Selkirk, 1984), and will be assumed here.

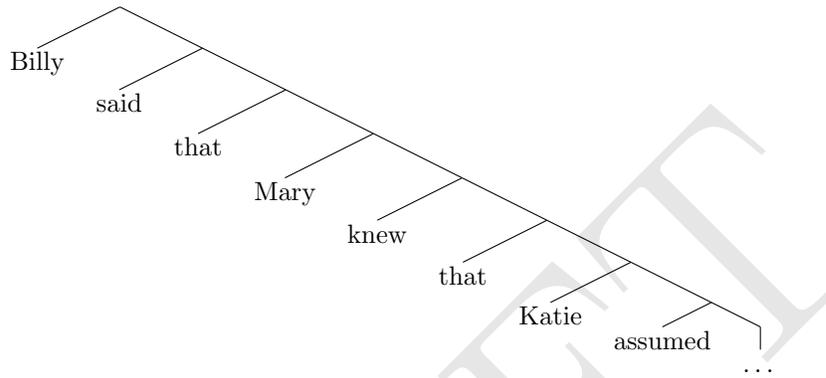


Figure 1: Syntactic structure: Binary and infinitely deep

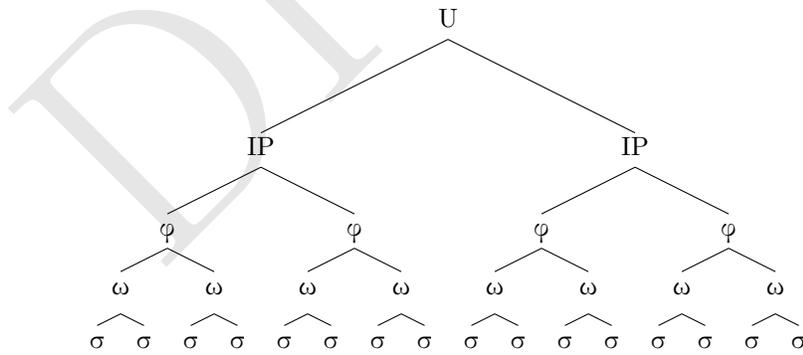


Figure 2: Phonological structure: Bushy and infinitely wide

What this means is that the language faculty has the onerous task of grafting the potentially profound dependencies of syntactic and semantic structure onto the phonological system, whose structure is defined by bushy and filled (i.e. non-covert) nodes of limited depth.<sup>1</sup> That said, there is an obvious and overwhelming tendency in language to align prosodic words/phrases with morphosyntactic words/phrases, a tendency fruitfully outlined in the constraint-based analysis of Selkirk (2011)'s *Match Theory*.

Pure matching of each syntactic phrase to each phonological phrase is, however, due to syntax's greater recursivity, impossible after a very limited amount of recursion, but we can say that languages conspire in various ways to alleviate matching violations. Here I'll argue that a number of mysterious syntactic phenomena can be understood as attempt to efficiently fit syntactic structure into phonological structure.

In Section 2.1, I will argue that the limitations on English DP ordering are the results of prosodic restrictions, later expanding these intuitions account for why many SOV languages nonetheless have post-verbal CPs (Section 2.2). I'll then theorize a general constraint on phonologically minor elements occupying the edges of prosodic phrases.

## 1.2 Phrasal Stress

In Section 3 I'll also suggest that the prosodic requirements of phonological structure can also come into play in 'narrowly' syntactic principles. Specifically, I'll address the growing body of literature seeking to address the External Projection Principle (Section 3.2) and that-trace effects as being phonologically motivated (Section 3.1) (Kandybowicz, 2006; Salzmann et al., 2011; Richards, 2012; McFadden and Sundaresan, 2015).

The argument is mainly based on the fact that these principles are alleviated by phonological changes and are conditioned by the boundaries of intonational phrases (IP). I'll argue that languages which show that-trace effects or the EPP have more highly ranked constraints that favor IPs with trochaic or at least initial stress, similarly linking this tendency to 'Wackernagalian' second-position phenomena.

## 2 Extraposition

### 2.1 English DP ordering

English is a familiar language with particularly liberal extraposition<sup>2</sup>, and as such is worth beginning with. English DP/NPs are usually realized in a canon-

---

<sup>1</sup>It should also be noted that non-phonological processing constraints may limit practical recursion in language as well. This is a complement to this argument here: theoretically infinite syntactic depth must, for various interface limitations, be variously 'flattened' for enunciation and comprehension.

<sup>2</sup>I should say that by extraposition, I mean this in the widest sense which encompasses all right-ward movement that breaks 'underlying' order with or without a resumptive *it* pronoun.

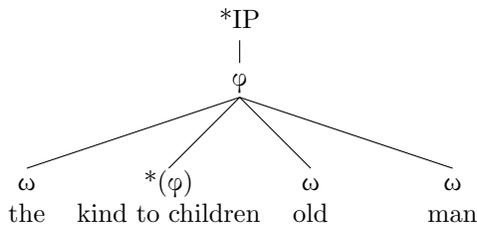


Figure 3: Prosodically faulty ‘underlying’ order in (3a)

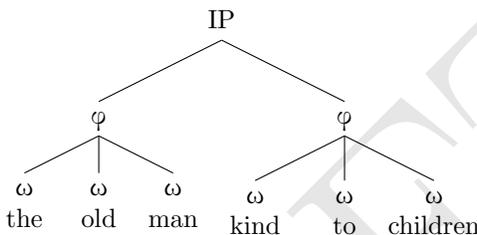


Figure 4: Prosodically well-formed derivation in (3b)

ical D-A-N surface order as below.

- (1) the kind old man
- (2) the surprised waitress

Interestingly, however, when adjectives take complements, thus becoming autonomous prosodic phrases, they may not remain in their ‘canonical’ positions, but are mandatorially extraposed to the right edge of the syntactic phrase.

- (3) a. \* the kind to children old man
- b. the old man kind to children
- (4) a. \* the surprised by the explosion waitress
- b. the waitress surprised by the explosion

This disparity makes sense given what we have said about grafting syntactic structure onto phonological structure. While the AP ‘kind to children’ is heavy enough to warrant a its own  $\varphi$ , it cannot project to this level phonologically as it is already part of another  $\varphi$  (illustrated in Figure 3).

However, the AP can surface as a distinct  $\varphi$  outside of the projection of the greater DP, thus we can enunciate the structure in Figure 4. Note also the fact that since the AP is a distinct  $\varphi$ , there is a brief prosodic break between ‘man’ and ‘kind’ reminiscent of an appositive break.

At that, if a native English speaker attempts to pronounce the base order of (3a), he will most likely force the adjectival phrase into one large prosodic word

(ω): “the *kind-to-children* old man.” This is consistent with English prosodic structure, albeit forced and dispreferred in natural language.

As an additional note, it’s also conceivable that English nouns modified by prepositional phrases like ‘the man in the hat’ are the result of extraposition from a structure like ‘the in the hat man.’ Such an analysis would easily bring these sentences under the adjuncts-as-specifiers analysis of the Cartographic Program, and could be matched with projections already posited for adjectives (Scott, 2002).

Regardless, the irregularity of English DP ordering can be construed as

## 2.2 CP Extraposition

A situation quite similar to English DP ordering is that of the systematic CP extraposition in many SOV languages. While some languages (Japanese, for example) are militantly SOV, a large number of SOV languages cannot accommodate preverbal CP complements, and thus they must occur verbal-finally. This is illustrated in German (SOV with V2 in main clauses) below.

- (5) a. Ich will **es** wissen.  
 I want it to know  
 I want to know it.
- b. Ich will wissen **was Liebe ist**.  
 I want to know what love is  
 I want to know what love is.
- (6) a. Uli hat **es** gesagt.  
 Uli has it said  
 Uli said it.
- b. Uli hat gesagt, **dass sie es sehen möchte**.  
 Uli has said that she it to see wants  
 Uli said she wants to see it.

We see the same alternation in Basque in (7) (Hualde and Ortiz de Urbina, 2003) and in Persian.

- (7) Udaltzainek ukatu dute [Rubioren bizkartzain zirela.]  
 policemen.ERG deny AUX [Rubio.GEN bodyguard were.C]  
 The town policemen have denied that they were Rubio’s bodyguards.

This motivation of this asymmetry can be compared to the aforementioned English examples, albeit on the level of the Intonational Phrase (IP). Full syntactic CPs generally necessitate independent IPs themselves, but if wedged between two other elements of the same syntactic level, already projecting to an IP, this is impossible given the Strict Layer Hypothesis.

This is again illustrated in Figure 5. *was Liebe ist* cannot properly project to an IP seeing that the elements to either side ideally project to the same

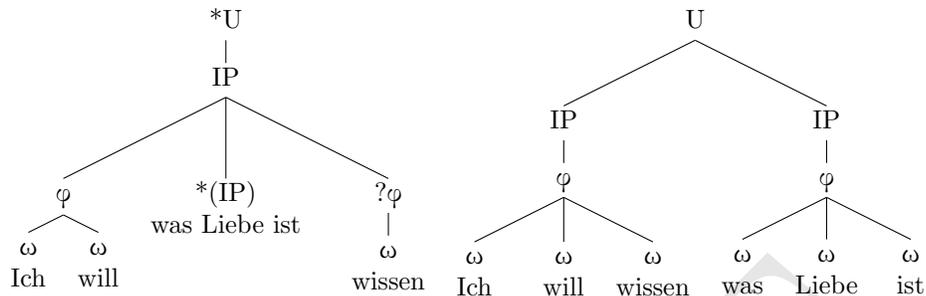


Figure 5: CP extraposition in German

IP. Additionally, one word: *wissen* is *stranded*, so to speak, projecting to a  $\varphi$  all alone, another potential suboptimality. By extraposition, both of these are solved as shown on the right.

### 2.3 Constraints against phonological stranding

This said, it should be noted that functionally, what is additionally ‘wrong’ with an utterance like (3a) is that if we map the AP ‘kind to children’ to a  $\varphi$  without movement, we have two phonological elements ‘stranded’ on either side: ‘the’ and ‘man’ which are intuitively too minuscule to project to their own  $\varphi$ , while they still syntactically belong to the same projection.

The same is true of extraposition of complements to V in SOV languages: if something in object position is heavy enough to warrant its own  $\varphi$ , the V is ‘stranded’ undesirably at the end of the utterance alone. By extraposing in each case, these stranded elements can be integrated into another  $\varphi$ , and the system can avoid awkwardly uncontentful  $\varphi$ ’s and economize on them generally (2 instead of 3).

Again, here I have motivated extraposition from a need for certain syntactic phrases to project to a particular layer of prosodic structure, in a way somewhat analogous to Selkirk (2011)’s *Match Theory*; this is not the only way of formalizing this. Yet even if we assume not prosodic matching nor even the Strict Layer Hypothesis, we can say that extraposition can be motivated from a prosodic need to economize on ‘branches’ and to minimize individual words projecting too high on the prosodic hierarchy without others.

## 3 The Trochaic Needs of Phrases

The phonological motivation of syntactic phenomena needn’t end there. All of this said, there seems to be good reason to assume that various languages favor particular prosodic structures within different layers of prosodic phrases. Particularly, there seems to be a general preferences for a strong trochaic start to phonological elements cross-linguistically. To conceptualize this, we can refer

to the STRONG START constraint of Selkirk (2011), reproduced below.

(8) **Strong Start** (Selkirk, 2011)

A prosodic constituent optimally begins with a leftmost daughter constituent which is not lower in the prosodic hierarchy than the constituent that immediately follow:

$$*(\pi_n \pi_{n+1} \dots)$$

STRONG START stipulates that the first element in any prosodic phrase take stronger stress than the second element. On the word level, this may be a syllable; on the phrase level, a word—regardless, STRONG START requires a kind of trochaic structure to prosodic phrases. A language with a highly ranked STRONG START constraint will

In the same vein, Fitzgerald (1994) provides a similar account of the phonologico-syntactic constraints of Tohono O’odham, a language which, despite extremely free word order, demands that the unemphasized sentential auxiliary be always the second element in the sentence.

Fitzgerald gives evidence for this auxiliary being ‘underlyingly’ an sentence initial particle which is shuffled to second position by a constraint that demands a trochaic sentence structure and thus, cannot stress the reduced auxiliary. This trochaic constraint, similar to STRONG START here, is also attested by disability to begin a Tohono O’odham clause with the stressless determiner, among other things. This kind of analysis and constraint could doubtlessly be expanded to other languages with heavy second-position alternations.

In the sections below, I’ll argue that STRONG START or similar constraints that demand trochaic structure at the beginning of phrases and can account for a variety of what might otherwise be thought of as narrowly syntactic alternations. Similar attempts have already been made fruitfully with similar phonological principles. An (2006), for example, posits the *Intonational Phrase Edge Generalization* to deal with the distribution and necessity of *that* in certain situations (as illustrated in (10)).

(9) **Intonational Phrase Edge Generalization (IPEG)** (An, 2006)

The edge of an I-phrase cannot be empty (where the notion of edge encompasses the specifier and the head of the relevant syntactic constituent).

- (10) a. I do know (that) he went to the store.  
b. \*(That) he went to the store I do know.

### 3.1 That-trace Effects

Now that-trace effects have traditionally be modeled as a somewhat arbitrary syntactic filter without much context or similarity in human grammar. To illustrate the classic problem, English seems to disbar the sequence of *that* followed by a subject extraction trace (shown in (11b)), while the equivalent sentence without *that* is acceptable (as in (11a)).

- (11) a. Who do you think gave Billy his black eye?  
b. \*Who do you think that *t* gave Billy his black eye?

As traditionally modeled, the idea had been that English instantiates a very specific syntactic filter simply weeds out the linear sequence *that-t*. Again, this seems like a very specific constraint to add to the possible toolbox of Universal Grammar.

### 3.1.1 That-trace effects are phonological in nature

A mounting testament of evidence, however, points to the fact that that-trace effects are, however, not a narrow syntactic hole, but are phonologically conditioned. Firstly, as Merchant (2001) has noted, ellipsis of the phonological form of a that-trace effect avoids the violation. Thus (12) remains grammatical, despite the fact that the syntax of the elided phrase would have to violate the that-trace effect filter.

- (12) John said that someone would write a new textbook, but I can't remember who (\*John said that *t* would write a new textbook).

At that, that-trace effects are also alleviated by the scrambling of adverbial phrases topicalized in clauses which would otherwise yield violations.

- (13) a. Who do you think that after all these months *t* would still want to go on the trip?  
b. Who did they say that for his whole life *t* wanted a new puppy?

Importantly Salzmann et al. (2011) realize explicitly that that-trace effects are not merely abstract collocations of *that* and a trace, but on the phonological level, are a linear collocation of *that* and a finite verb. The flexibility of German scrambling can collocate a variety of elements with *dass*, but only the main verb produces the stark unacceptability.

### 3.1.2 The Motivation of *that-trace* Phenomena

The importance of all of this for our purposes is that the phonological constraints can interact with functional tendencies in language in a systematic way. Crucially, Gundel (1988) recognizes that languages tend to align prosodic and sentential stress to topicalized and focused entities, which occur in cross-linguistically similar locales; these elements are routinely nouns or topical PPs. At that, Bolinger (1954) makes a similar descriptive statement of Spanish, where the surface realization of nominals is often a function of the wider prosodic structure of a sentence; that is, a focused noun will tend to fall where the natural stress of an utterance would otherwise be. In formal literature as well (Rizzi, 1997) (and intuitively), it should be noted that the elements which are focused or topicalized in language are routinely nominals, not verb heads.

Drawing this all together, we can say that that-trace effects actually fall out merely from STRONG START. Assuming that *that* produces a new intonational

phrase boundary after it, the following element which begins the IP must be prosodically more prominent than its successor. However main verbs usually have *no* focus or topic quality, they cannot serve this purpose, and thus cannot begin intonational phrases.

Thus a language with a high ranking on a STRONG START-like constraint will rule out verb-initial clauses in the same way that languages with productive second-position phenomena (like Tohono O’odham mentioned above) will rule out clitic-initial clauses. Here I will be assuming that STRONG START is a highly ranked constraint for English IPs (while in the canonically VSO languages like Irish, it is ranked lower).

As one should expect, in positions where verb heads happen *to* serve as objects of contrastive focus, that-trace effects are somewhat alleviated as in (14) (Kandybowicz, 2006).

- (14) % Who did you say that WROTE Barriers today?

English CPs prefaced by *that* must project to full IPs, thus potentially paving the way for the violations of STRONG START. On the other hand, English CPs without *that* are enunciated as a single IP phrase, thus the lack of a preverbal element is not prosodically dispreferred.

Interestingly, if a that-trace effect-violating phrase (like 11b)) is read quickly enough as one IP, the violation is alleviated or annulled. A similar fact is noted by Kandybowicz (2006), saying that that-trace effect-violating sequences become acceptable or improved when either the C or following verb is prosodically collapsed onto one another.

- (15) a. Who do you hope \*for/%fer to win?  
 b. % Who do you suppose that’ll leave early?  
 c. % The author that the editor predicted that’d be adored. . .

In examples like (15), we can say that the cliticization nudges the reader to read the sentences as a single, uninterrupted IP, thus avoiding violations of STRONG START. These can be compared to the above example of forcing strings like ‘the kind to children man’ into acceptable English prosody (such that we have ‘kind-to-children’ realized as one prosodic word).

Additionally, in the same way that complement CPs with *that* project to IPs and CPs without do not, we can see that *restrictive* relative clauses like (16a) do not produce that-trace effects while their near cousins, *non-restrictive* relative clauses, which do have distinct intonational boundaries (comma intonation) do (16b).

- (16) a. the fisherman that *t* liked to play dice  
 b. \* the fisherman, that *t* liked to play dice

The required IP boundary in a phrase with a non-restrictive relative clause puts the syntactic clause up to STRONG START on IPs. Again, where there is no IP boundary, like in (16a) or in a complement CP without *that*, STRONG START does not apply.

## 3.2 The EPP

### 3.2.1 The EPP is phonological in nature

The EPP can be modeled in much of the same way, indeed the classic assumption had been that the EPP was *the same* or at least similarly parameterized as that-trace effects.

Again, STRONG START requires a IP to begin with an element higher on the prosodic hierarchy than the following element, and given the lower prominence of main verbs, this amounts to a ban of IP-initial verbs. Intuitively, even in stylistic scrambling in English, verb initial phrases seem highly cacophonous (although the distinct question intonation makes can permit verb-initiality).

- (17)
- a. I like peanuts.
  - b. Peanuts I like.
  - c. Peanuts like I.
  - d. ?? Like I peanuts.
  - e. ?? Like peanuts I.
  - f. Like I peanuts?

Even in English clauses with scrambled order, declarative intonation causes an activation of STRONG START.

I was going to write more here, but then I didn't.

### 3.2.2 An additional little addendum on the EPP

Bear in mind the esoteric nature of the EPP in comparison with the framework outlined in Section 1. There is a general sense in which the EPP makes syntactic structure more amenable to the phonological system: that's to say, it takes a potentially very heavy verb phrase, containing a head and potentially three nominal arguments and extracts one to a new domain. In this sense, it does the division that prosodic matching would need. In this sense, 'deep' syntactic structure might be constructed in a way very tight and unmalleable to linearization or phonological chunking, but the EPP surgically removes an element which allows a syntactic division less complex in depth thus facilitating prosodic grafting.

This is partially a functional explanation. Still, there is a general problem stewing in all languages: that syntactic and prosodic structure do not match, and different languages may have functionally developed different strategies for relieving it: the EPP, overt movement, discourse-configurationality, scrambling, perhaps even word order differences etc. All of these generally work together to spread Externally Merged structure about in the derivation, leaving smaller chunks which can be more easily phonologized into the finitely deep structure of phonology.

## 4 Closing Notes: *Syntax* is the Interface

There has been a decent amount of existential angst surrounding the issues of ‘the syntax-phonology interface.’ Indeed, in dealing with data similar to those above, McFadden and Sundaresan (2015) and Richards (2010) grapple with the apparently ‘counter-cyclicity’ of syntactic alternations which are drawn out from phonological constraints. Indeed, in the classic formulation of generative grammar, the idea had been that a syntactic engine produced a set of valid strings which on one hand changed by transformations and later enunciated by the phonological system, while on the other hand, meaning is computed from the syntactic deep structure.

Minimalism retains a simplified version of this, in which the ‘transformations’ are a part of the gradual and uniform syntactic derivation itself (in the form of Internal Merge) and phonological form and the semantic content are derived from the output.

The problem for the data here is that phonology is *downstream* from syntax. The phonological system may have its own obvious constraints, but if it is fed a full syntactic string to enunciate, it can’t go back in the system and condition further changes. This has been dealt with, like too much else, with “features,” which are a mere formalization of a description.<sup>3</sup> At that, features end up creating a kind of redundancy in the language faculty where two different modules share the same restrictions (similar to the redundancy of C selection in the semantics realm).

More clarity would be induced by stepping a step back. As mentioned in Section 1, the language production system must take a base structure of meaning/semantics (syntactic ‘deep structure’) and graft it into the phonological system. The thing we call *syntax* is an epiphenomenon of the interface between the phonological and semantic systems, a messy spawn of two fundamentally different systems.

Of course we can imagine a supreme elegance and consistency in phonology in the abstract, and a supreme elegance in semantic structure as well, but when the two are shoe-horned together, you get nasty things like extraposition and wh- movement and things that don’t make any sense by themselves. Thus syntax is not the core of the language faculty which produces outputs for the interfaces of sound and meaning, but it itself is an interface between the two.

### 4.1 Towards a Plausible Mechanism

How precisely the language faculty grafts semantic structure into prosodically viable phrases (language production) and how it parses prosodic phrases into semantic structure (comprehension) would still be unclear. But as with any

---

<sup>3</sup>What I mean by this is that a pre-generative description of an alternation, such as, in Z condition, X moves to Y is no less communicative than a ‘formal’ analysis saying that X moves to Y to check Z feature. In fact, the latter explanation simply adds on theoretical machinery without explanation of explanation.

system we don't understand anything about in linguistics, we can at least tentatively model it with Optimality Theory or another constraint based system, which is not to be summed to the exclusion of a more psychologically plausible theory. Here I've suggested some constraints that weed sentences know to be suboptimal.

English DP ordering, as well as CPs in SOV languages and second-position phenomena are all plausibly phonologically motivated, as is the unacceptability of subjectless or that-trace effect-violating clauses.

Now there have been quite interesting attempts to motivate the EPP and other principles mentioned here on semantic or pragmatic grounds (Butler, 2004). It should be said that this explanation is not necessarily to the exclusion of these analyses. Indeed, my implicit argument here would be that a conspiracy of factors reinforce each other to yield certain structures or epiphenomena like the EPP.

Indeed, if it is the case that there is a full functional structure of which encodes extra semantic or pragmatic meaning (à la Rizzi (1997)), we can still say that whether a projection 'moves' overtly to one of these nodes which affects them is a function of prosodic constraints. This can be compared to Richards (2010)'s model of phonologically-motivated *wh*-movement, where what determines where a *wh*-word will be pronounced is merely how great the number of prosodic phrase boundaries there are between the *wh*-word and C. For our purposes, what determines which 'link' in a syntactic 'chain' is pronounced might be better modeled as a function of phonological or prosodic constraints of a language.

## References

- An, D.-H. (2006). Clauses in noncanonical positions at the syntax-phonology interface. *Syntax*, 10.1:38–79.
- Bolinger, D. L. (1954). English prosodic stress and Spanish sentence order. *Hispania*, 37.2:152–156.
- Butler, J. (2004). On having arguments and agreeing: Semantic EPP. *York Papers in Linguistics*, 2.1.
- Fitzgerald, C. (1994). Prosody drives the syntax: O'dham rhythm. *Proceedings of BLS 20*.
- Gundel, J. K. (1988). Universals of topic-comment structure. In *Studies in Syntactic Typology*. John Benjamins Publishing Company.
- Hualde, J. I. and Ortiz de Urbina, J. (2003). *A Grammar of Basque*. Mouton de Gruyter.
- Kandybowicz, J. (2006). Comp-trace effects explained away. *Proceedings of the 25th West Coast Conference on Formal Linguistics*.

- McFadden, T. and Sundaresan, S. (2015). Towards resolving the countercyclic-ity of the EPP.
- Merchant, J. (2001). *The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis*. Oxford University Press.
- Richards, N. (2010). *Uttering Trees*. MIT Press.
- Richards, N. (2012). Generalizing a metrical EPP. In *Workshop on Locality and Directionality at the Morphosyntax-Phonology Interface*.
- Rizzi, L. (1997). The fine structure of the left periphery. In Haegeman, L., editor, *Elements of Grammar*. Kluwer.
- Salzmann, M., Häussler, J., Bader, M., and Bayer, J. (2011). That-trace effects without traces: An experimental investigation. *Proceedings from the 42nd Annual Meeting of the North East Linguistics Society*, 2.
- Scott, G.-J. (2002). Stacked adjectival modification and the structure of nominal phrases. In Cinque, G., editor, *Functional Structure in DP and IP*, pages 91–120. Oxford University Press.
- Selkirk, E. (2011). The syntax-phonology interface. In Goldsmith, J., Riggle, J., and Yu, A. C. L., editors, *The Handbook of Phonological Theory*. Blackwell.
- Selkirk, E. O. (1984). *Phonology and Syntax: The Relation Between Sound and Structure*. MIT Press.
- van der Hulst, H. (2005). Phonology is the same. In Broekhuis, H., Corver, N., Huybregts, R., Kleinhenz, U., and Köster, J., editors, *Organizing Grammar: Linguistic studies in honor of Henk van Riemsdijk*, pages 252–262. Mouton de Gruyter.