Kruger, William W. 2023. Prosodic conditioning of the null complementizer in English. *Glossa: a journal of general linguistics* 8(1). pp. 1–35. DOI: https://doi.org/10.16995/glossa.8860

Open Library of Humanities

Prosodic conditioning of the null complementizer in English

William W. Kruger, Arizona State University, US, wkruger@asu.edu

This paper presents a novel account of the distribution of the null complementizer (C) in English, arguing that the licensing of null C is a case of prosodically conditioned allomorphy (Carstairs 1988; 1990; Paster 2006; Inkelas 2014). The account expands on prior proposals (Bošković & Lasnik 2003; An 2007a; b), observing that environments which prohibit null C all show an obligatory Intonational Phrase (IP) boundary preceding C. In contrast, an obligatory IP boundary does not appear in environments which allow null C. I argue that null C is an allomorph associated with a specific prosodic environment: "medial" position within IP, (...C...). Overt C, on the other hand can appear in "initial" position within IP, (C...). This synchronic account is supported by a diachronic account, tracing how null C arose from phonologically weak overt forms of C (/ðæt/ → /(ð)æt/, /(ð)ət/, /-t/ → ∅). Crucially, the distribution of these weak forms is itself prosodically conditioned by principles which allow phonological weakening in medial prosodic positions but prevent weakening in initial position, a result of "domain-initial strengthening" (Pierrehumbert & Talkin 1992; Fougeron & Keating 1997; Beckman 1998; Keating et al. 2004; White & Turk 2010).

Glossa: a journal of general linguistics is a peer-reviewed open access journal published by the Open Library of Humanities. © 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See http://creativecommons.org/licenses/by/4.0/.
OPEN ACCESS

1 Introduction

The distribution of the null finite complementizer (C) in English has been a longstanding puzzle. A small selection of the basic facts are as follows:

(1) a. I believe **{that, Ø}** Sue likes linguistics.
b. **{That, *Ø}** the teacher was lying was hardly obvious.
c. I believe, based on intuition, **{that, *Ø}** Sue likes linguistics.

Numerous attempts have been made to capture why both overt C (*that*) and null C (Ø) are allowed in (1a) but null C is disallowed in (1b) and (1c). These range from purely syntactic accounts (Stowell 1981; Webelhuth 1992; Pesetsky 1995; Pesetsky & Torrego 2001) to morphosyntactic and linearization-based accounts (Bošković & Lasnik 2003; Bošković 2005; Lohndal & Samuels 2013) to prosody-based accounts (An 2007a; b), as well as an array of variationist, usage-based, and frequency-based accounts (see overviews in Shank et al. 2016; Liang et al. 2021). These many different approaches show varying degrees of success in capturing the full range of data and in demonstrating sufficient explanatory power. For our purposes here, it suffices to say that despite significant advancement in our understanding, the phenomenon of null C distribution remains an open question.

In this paper, I will pursue a line of thinking not yet represented in the literature cited above: null C is a *prosodically conditioned allomorph* of the finite complementizer morpheme. Prosodically conditioned allomorphy is a subcase of phonologically conditioned allomorphy (Carstairs 1988; 1990; Paster 2006; Inkelas 2014) where a given allomorph is licensed by its local prosodic environment. In short, the proposal of this paper is that the null allomorph can only be selected when C occupies a *medial position within a prosodic constituent*, i.e. non-adjacent to a prosodic boundary. This idea builds upon a set of observations about the range of environments which prohibit null C, going back at least to Bošković & Lasnik (2003), later taken up by An (2007a; b). The observation is that contexts in which null C is prohibited all show the obligatory presence of an *Intonational Phrase* (IP) boundary immediately preceding C. In other words, null C is prohibited when it would otherwise occupy an *initial position within a prosodic constituent*. This can be conveniently seen in (1b) above, where C heads a Clausal Subject (*That the teacher was lying*) and is adjacent to the initial boundary of the sentence itself, and in (1c), where a Parenthetical breaks up the prosodic phrasing of the sentence, represented by comma punctuation. The full range of contexts which prohibit null C that will be addressed in this paper to support the prosodic generalization is as follows: Clausal Subjects, Topicalized Clauses, Extraposition, Parentheticals, Non-Bridge Verbs, Noun Complement Clauses, Right-Node-Raising, Gapping, and Pseudoclefts.

The proposal outlined so far is a synchronic account of the facts, but it is crucial to consider diachrony as well. To that end, I will also address how the prosodically conditioned allomorph Ø can be understood as deriving from progressively weakened forms of overt C in the history of

English (e.g., /ðæt/ > /ðət/, /(ə)t/ > /Ø/)—forms which are themselves conditioned by prosodic boundaries. This will require a broader discussion of prosodic constraints on the phonological reduction of English function words (prepositions, auxiliaries, determiners, etc.). In particular, I rely on facts about the application of vowel reduction and deletion of word-initial segments. (2a) below shows that deletion of word-initial /h/ on forms of the auxiliary verb *have* is allowed in medial position, while (2b-c) show deletion prohibited when preceded by a prosodic boundary, matching the distribution of null C in (1a-c) above.

(2) a. Sue **has [æz]** driven to London.
b. **Has [hæz]** Sue driven to London?
c. Sue, if I'm not mistaken, **has [hæz]** driven to London before.

The segmental strengthening in (2b) and (2c) is referred to as "Domain-Initial Strengthening (DIS)", a well-established crosslinguistic and diachronic effect whereby stronger articulation of segments is maintained in the initial position of prosodic constituents (Lass & Anderson 1975; Escure 1977; Pierrehumbert & Talkin 1992; Fougeron & Keating 1997; Beckman 1998; Cho & Keating 2001; Fougeron 2001; Keating et al. 2004; White & Turk 2010). For convenience, I will refer to the inverse weakening effect in (2a) as "Domain-Internal Weakening (DIW)".

The further claim of this paper is that, in cases where a given morpheme develops allomorphs over time that correlate with prosodic position (i.e. the emergence of prosodically conditioned allomorphy), "strong" forms of allomorphs will be associated with initial prosodic positions, while "weak" forms will be associated with medial prosodic positions. In the case of the finite complementizer, over time, DIW causes C to undergo phonological reduction, eventually to zero in the late Middle English period, and the zero-form emerges as a distinct null allomorph correlating with medial prosodic positions. In contrast, DIS prevents reduction in domain-initial position, and so the "strong" form of the complementizer (i.e. *that*) is preserved in initial position. Note that selection of null C is a case of optional allomorphy, since overt C can also still be selected in medial position. DIW simply yields an extra optional choice for the form of the morpheme, one that has been diachronically conditioned (and subsequently synchronically acquired by learners) to correlate with medial prosodic positions.

The paper is organized as follows: In Section 2, I present the full dataset on null C distribution along with the current state of our understanding of the prosodic properties of the relevant sentences. In Section 3, I briefly address and critique prior accounts of this data, focusing specifically on the prosodic account of An (2007a; b). Section 4 lays out the main proposals of the paper (briefly sketched above), split into a synchronic component, whereby null C is characterized as a prosodically conditioned allomorph, and a diachronic component explaining the historical development of null C and the origins of prosodic conditioning. Section 5 discusses some interesting consequences of the account, while Section 6 concludes.

2 Data on null C distribution

This section presents the facts on null C distribution, starting with sentences that allow null C and then moving to cases where null C is prohibited. Anticipating the main proposal of the paper (that null C is a prosodically conditioned allomorph), I will also include information on the prosodic form of these sentences, focusing on aspects of prosodic phrasing and the assignment of Intonational Phrases in particular.

The English finite complementizer may optionally be phonologically overt (*that*) or phonologically null (Ø) in the contexts shown below, adapting data and judgements from Aoun et al. (1987) and An (2007b), which trace back to work by Rosenbaum (1965), Perlmutter (1968), Postal (1974), Stowell (1981), Hornstein & Lightfoot (1991), and Pesetsky (1995).

(3) a. I believe **{that, Ø}** Sue likes linguistics.
b. It was apparent **{that, Ø}** Kay left.
c. The book **{that, Ø}** Kay wrote arrived. (An 2007b: 39, Aoun et al. 1987: 540)

Each of these cases shows an embedded finite clause immediately following a lexical word, such as a verb (3a), adjective (3b), or noun (3c). In (3a-b), the syntactic structures for these sentences are typically characterized as involving local complementation or head-argument structures, while (3c) is a restrictive object relative clause, the precise syntax of which is beyond the scope of this paper. These facts and related ones are routinely cited in syntactic literature and are easy to confirm (Perlmutter 1968: 215; Stowell 1981: 409–412; Quirk et al. 1985: 1222ff; Aoun et al. 1987: 540–545; see also Mindt 2011: 149 et seq. for many examples with adjectives). Going forward, I will primarily focus on (3a) as the most commonly cited case where null C is licensed.

Turning to the prosodic features of these sentences, an important first step is to outline properties of intonational phrasing in English. Briefly, an *Intonation Phrase* (IP) is defined as a prosodic unit which is flanked by pauses and/or boundary tones and which contains a number of pitch accents. It is commonly assumed that the syntactic level of the "clause" is obligatorily parsed as an IP, along with a number of other syntactic items, including parentheticals, extraposed and dislocated constituents, tag questions, and vocatives (Selkirk 1978; 1984; 1996; Nespor & Vogel 1986). For the sentences in (3) above, however, the prosodic feature of note is that both the main clause and the embedded clause can optionally be parsed into separate IPs or *together into a single IP* (Emonds 1970; Downing 1970; Selkirk 1978; 1984; 2005; 2011; Pierrehumbert 1980; Nespor and Vogel 1986; Ladd 1986; 1996; Truckenbrodt 2005; Downing 2011). The sentences in (4) illustrate these two options (from Nespor & Vogel 1986: 199):

(4) a. ($_{IP}$ Our next door neighbor truly believes) ($_{IP}$ that black cats bring bad luck)
b. ($_{IP}$ Our next door neighbor truly believes that black cats bring bad luck)

Note that the pronunciation of the first IP in (4a) will show a "continuation rise" on *believes* (a pitch accent on the first syllable with a drop-off in the second syllable, followed by a pitch-rise and a short pause), signaling a juncture between IPs.[1] The most important idea going forward is that all three cases in (3) are known to allow the "integrated" IP-parsing shown in (4b), whereas the bulk of evidence suggests that the rest of the sentences discussed below—many of which contain diverse syntactic constructions with varying rules and derivations—do not allow integration between the main clause and embedded clause. As a result, the common thread amongst all of the sentence-types below is that they involve a configuration where the complementizer is obligatorily adjacent to an initial IP-boundary.

Let's now turn to cases where overt C is allowed, but null C is disallowed. As will be seen, the number of environments in this category far outnumber the cases where null C is allowed.

(5) a. **{That,** ***Ø}** the teacher was lying was hardly obvious. Clausal Subject
b. **{That,** ***Ø}** John likes Mary, Jane didn't believe. Topicalized Clause
(Rosenbaum 1965: 69; Stowell 1981: 396–397; Bošković & Lasnik 2003: 529)

(5) shows two cases where C is positioned at the beginning of the main sentence, a Clausal Subject in (5a) and a Topicalized Clause in (5b), and null C is strongly prohibited. In fact, the prohibition on null C in sentence-initial position is likely one of the more consistent judgements in the literature. There is not much to say here in terms of prosody, other than to note that (5a) can be straightforwardly parsed into at least two IPs: ($_{IP}$ **That** the teacher was lying) ($_{IP}$ was hardly obvious). It is unclear if an "integrated" IP-phrasing is available for (5a), although this is immaterial for our purposes here. The same is true for (5b), although it is known that topicalized constituents are typically obligatorily parsed into separate IPs, maintaining a Topic-Comment information structure (Downing 1970: 41–45), so maintenance of two IPs is to be expected. The upshot is that, in both cases, C is always obligatorily preceded by an IP-boundary.

(6) shows two cases where adjacency between an embedded clause and a preceding lexical word (the verb *believe*, in this case) is disrupted: Extraposition in (6a), a process which shifts the clause rightward (Aoun et al. 1987: 540–541), and insertion of a Parenthetical between verb and embedded clause in (6b). In both cases, null C is judged to be disallowed.

(6) a. I believe very strongly **{that,** ***Ø}** Sue likes linguistics. Extraposition
b. I believe, based on intuition, **{that,** ***Ø}** Sue likes linguistics. Parenthetical
(Quirk et al. 1972: 735; 1982: 276; An 2007a: 39)

[1] Note that these options for IP-parsing are by no means unique to English, and have also been observed for Xhosa (Jokweni 1995), Turkish (Kan 2009), Catalan (Feldhausen 2010), Swedish (Myrberg 2010; 2013), German (Truckenbrodt 2015), and Hungarian and Bàsàá (Hamlaoui & Szendrői 2017). Some languages have also been identified as exceptions, assigning IP-boundaries to all clauses regardless of their embedded status; see Ishihara (2014) on Japanese, Pak (2008) for Luganda.

Prosodically, both of these cases are fairly well-described. It is routinely reported that extraposed objects are obligatorily parsed into a separate IP from the IP of the main clause, with a clear prosodic break between them (Nespor & Vogel 1986: 188; Hale & Selkirk 1987; Zec & Inkelas 1990: 376–377; Truckenbrodt 1995; 1999; Bošković 2001; 2005; An 2007a; b). Thus, (6a) must be parsed as ($_{IP}$ Sue said yesterday) ($_{IP}$ **that** Kurt had left), with no discernable option for integrated IP-parsing. The case of parentheticals is similar: it is well-established that parenthetical phrases "break up" the prosodic and syntactic material into which they are inserted, and as noted above, parentheticals are a part of the class of syntactic objects which are known to be obligatorily parsed as separate IPs, with attendant pauses flanking the parenthetical and corresponding continuation rise (see Dehé 2009: 571 and citations therein; see also Nespor & Vogel 1986: 188–190; Selkirk 1984: 295–296; and discussion in Gussenhoven 2004; Astruc 2005; Dehé & Kavalova 2007; and Kaltenböck 2007). (6b) therefore must be parsed as ($_{IP}$ I believe) ($_{IP}$ based on intuition) ($_{IP}$ **that** Sue likes linguistics). Note that both of these cases result in configurations where C is obligatorily preceded by an IP-boundary.

Given the dataset so far, it might be tempting to explain licensing of null C by appealing roughly to the pattern of adjacency between the C of the embedded clause and the preceding lexical word (an idea that is indeed explored by Pesetsky 1995 and Bošković & Lasnik 2003). Unfortunately, however, not all lexical words license null C in a following embedded clause. To start, (7) shows null C disallowed after "non-bridge verbs", a class containing "manner of speaking" verbs (*quip, whisper*), subject experiencer verbs (*love, hate*), and "factive" verbs (*regret, resent*). The identification of these types of verbs goes back to Ross (1967: 252), Zwicky (1971), and Erteschik-Shir (1973), see also Stowell (1981: 397).

(7) Fay quipped **{that,** ***Ø}** Kay left. Non-Bridge Verb
(Aoun et al. 1987: 544)

The intonational phrasing of sentences with verbs like *quip* is addressed by Franks (2000; 2005), citing the judgements in (8)-(9) below. (8a-b) shows that an IP-boundary may precede, *but not follow*, the C of an embedded clause after a non-bridge verb like *quip*, while, interestingly, (9a-b) shows that a boundary may optionally *precede or follow* C after a verb like *say*.

(8) a. ($_{IP}$ Bill quipped) ($_{IP}$ **that** he saw a ghost)
b. *($_{IP}$ Bill quipped **that**) ($_{IP}$ he saw a ghost)

(9) a. ($_{IP}$ Bill said) ($_{IP}$ **that** he saw a ghost)
b. ($_{IP}$ Bill said **that**) ($_{IP}$ he saw a ghost) (Franks 2005: 18–21)

Franks also judges that, for a sentence with *quip*, "having a break before *that* is highly preferred to pronouncing these sentences as a single Utterance" (2005: 19). In other words, there is a

preference to maintain separate IPs for both main clause and embedded clause in sentences with non-bridge verbs. It is also relevant, according to my judgement, that sentences with *quip* vs. *say* differ in the appearance of a phrasal pitch accent on the verb. Accent is obligatory on *quip*, likely representing a boundary tone, but only optional on *say*, correlating with the integrated IP-phrasing discussed in (4b) above. The presence of a pitch accent has also been independently noted for factive verbs like *resent* and *regret* (also non-bridge verbs), which always receive accent when followed by an embedded clause (Kallulli 2006: 214; de Cuba & Urogdi 2010: 47–48).

Another instance of a lexical word not licensing null C is shown in (10), a "Noun Complement Clause (NCC)" construction where an embedded clause expresses the informational content of a preceding noun (*claim, story,* etc.):

(10) I distrust the claim **{that, *Ø}** Kay had left the party. Noun Complement Clause
(An 2007a: 40; Stowell 1981: 397–398)

Intonational phrasing of NCCs is analyzed by An (2007b), who first notes that NCCs have been traditionally analyzed as appositives—modifiers which are standardly assumed to parse into separate IPs, similar to parentheticals (Stowell 1981). However, there is little overt discussion in prior literature of the intonational phrasing of English NCCs. Instead, An (2007b) presents evidence from NCCs in Serbo-Croatian, Tagalog, and Brazilian Portuguese, showing that NCCs are standardly parsed into separate IPs in these languages, and by assumption, across languages. The interested reader is directed to An (2007b: 39–42) for overall discussion, as well as Bošković (2001) for Serbo-Croatian, Richards (1999) for Tagalog, and Guimarães (1998) for Brazilian Portuguese. Adding to this, it is my judgement that the same pattern identified by Franks (2000; 2005) for non-bridge verbs also applies straightforwardly to NCCs. In other words, an IP-boundary may precede (=11a), *but not follow* (=11b), the complementizer of a NCC, and a prosodic break after the head noun is more preferred to parsing the entire sequence as a single unit (=11c).

(11) a. ($_{IP}$ I distrust the claim) ($_{IP}$ **that** Sue had left the party)
b. *($_{IP}$ I distrust the claim **that**) ($_{IP}$ Sue had left the party)
c. ?*($_{IP}$ I distrust the claim **that** Sue had left the party)

Taken together, the facts on Non-Bridge Verbs and NCCs strongly suggest that a common feature of these constructions is the obligatory presence of an IP-boundary preceding C, a consequence of the apparent prohibition on integrated IP-parsing for these sentence-types.

The final set of sentences which do not license null C also mostly involve adjacency between C and a preceding lexical word. However, in each case, some additional operation has applied to the sentence, such as displacement, coordination, and/or ellipsis, depending on the preferred analysis. (12a) contains an example of Right-Node-Raising (RNR), consisting of an embedded clause placed in a coordination structure with two clause-taking verbs (Postal 1974; Aoun et

al. 1987). (12b) shows (Verb-)Gapping, where two embedded clauses are coordinated and the verb of the second conjunct has undergone ellipsis under identification with the verb of the first conjunct (Kuno 1976; Sag 1980; Aoun et al. 1987; Johnson 1997). (12c) shows a Psuedocleft, a construction consisting of a copula joining together an embedded *wh*-clause in the preverbal subject position and (for our purposes) an embedded clause in postverbal position (Postal 1974; Pesetsky 1995; Den Dikken et al. 2000; Den Dikken 2006; Herment & Leonarduzzi 2012).

(12) a. They suspected, and we believed, **{that, *Ø}** Peter would visit the hospital. RNR
b. Mary believed Peter finished school, and Bill, **{that, *Ø}** Peter got a job. Gapping
c. What the students believe is **{that, *Ø}** they will pass the exam. Pseudocleft
(Bošković & Lasnik 2003: 529)

Starting with RNR, the prosodic facts are described by Abbot (1976) and Swingle (1993), who observe that each component of an RNR construction (the two conjuncts and the rightward element) must be parsed into separate IPs, indicated by the typical comma punctuation in (12a) above (see also An 2007b: 47–48 for additional discussion): ($_{IP}$ They suspected) ($_{IP}$ and we believed) ($_{IP}$ **that** Peter would visit the hospital). Selkirk (2002) also finds a clear preference amongst speakers for a prosodic break after each of the coordinated verbs under RNR, which are themselves pronounced with pitch accents indicating narrow focus (see also Hartmann 2000; Féry & Hartmann 2005; Ha 2008 for similar observations). For Gapping, Selkirk (2002; 2005) again identifies a clear prosodic break after the gapped word, although the boundary may not be as strong as an IP-boundary (Selkirk defines it as a "Major Phrase" boundary): ($_{IP}$ Mary believed Peter finished school) ($_{IP}$ and Bill) ($_{?IP}$ **that** Peter got a job).

Intonational properties of pseudoclefts have also been identified in past work. Both Collins (2006: 1713) and Herment & Leonarduzzi (2012: 714) find that speakers parse pseudoclefts into two or more IPs in the majority of cases. Higgins (1979: 8) and Declerck (1984: 253; 1988: 5) observe a consistent IP-boundary after the copula of a pseudocleft, referred to as "colon intonation", shown in (13a) below (see Frascarelli & Hinterhölzl 2007; Frascarelli & Ramaglia 2009 for similar effects in Italian and German). This intonational pattern is associated with a "specificational" interpretation of pseudoclefts, where the postverbal constituent acts as a "value" for the "variable" expressed in the pre-verbal *wh*-clause (Higgins 1976: 95, Declerck 1988: 6–7; Hartmann 2006: 8). However, findings by Collins (1991: 81), Herment & Leonarduzzi (2012: 716), and Van Luven (2018: 62–65) complicate this picture somewhat, showing that pseudoclefts can sometimes manifest with an IP-boundary *before* the copula as well (=13b).

(13) a. ($_{IP}$ What the students believe is) ($_{IP}$ **that** they will pass the exam)
b. ($_{IP}$ What the students believe) ($_{IP}$ is **that** they will pass the exam)

My impression is that the IP-parsing in (13b) is indeed acceptable, and as such, pseudoclefts present the only (partial) counterexample to the otherwise consistent prosodic generalization

outlined so far: that null C is prohibited in sentences which show an obligatory IP-boundary before C. Note, however, that Kearns (2007) actually finds evidence that null C is indeed allowable in cleft sentences, contrary to the overall judgements cited in other literature. It is possible that the variation in IP-phrasing discussed above may be responsible for these inconsistencies, see below for fuller discussion of data from Kearns (2007).

This concludes the overview of environments which allow and prohibit null C. To recap, null C is allowed in sentences with local complementation or a head-argument relation between an embedded clause and a preceding lexical word (=3). In contrast, null C is prohibited with Clausal Subjects and Topicalized Clauses (5), with Extraposition and Parentheticals (6), with Non-Bridge Verbs (7), Noun Complement Clauses (10), and with Right-Node-Raising, Gapping, and Pseudoclefts (12).

The data presented here so far relies on the wealth of acceptability judgements reported in previous literature. However, it should be noted that the characterization of different cases as either purely licensing or purely prohibiting null C is likely an oversimplification. Kearns (2007: 312) presents a corpus study of written English showing significant variation in the distribution of null C in many of the above constructions, and these results can provide quantitative grounding for acceptability judgements. Kearns finds, for example, that null C appears 48–55% of the time with clausal complements to verbs and adjectives, largely as expected given the judgements in (3) above. More unexpected, however, is the finding that null C also appears in "prohibiting" environments, such as with Extraposition (31% of sentences), environments with *it*-subjects (20%) (a variant of Extraposition), and some cleft constructions (~9%). Even so, note that frequencies of null C in these environments are significantly reduced, compared to the standard licensing environment. This is important because it suggests that there is something about the environments in (5)–(7), (10), and (12) which creates a *dispreference* for null C, even if null C is not fully banned in every case. Any model which attempts to explain the distribution of null C should, therefore, be flexible enough to allow for variation while also explaining which constructions show a dispreference for null C and why.

I believe the account based on prosodically conditioned allomorphy articulated in Section 4 will fulfill this aspiration, since it does not attribute the distribution of null C to categorical syntactic, morphological, or semantic factors, which might predict more robust splits in the acceptability of sentences. An account of null C based in prosody has the advantage of locating the licensing principles in a module of language (prosody) which is at the intersection of various systems, including morphosyntax and semantics to be sure, but also allowing room for factors like variable prosodic phrasing, which is influenced by discourse context, information structure, and the size and length of syntactic units. Before continuing to the main account, Section 3 briefly addresses some prior proposals for null C distribution, focusing specifically on the syntax-prosody mapping account by An (2007a; b).

3 Prior accounts of null C

Numerous attempts have been made to capture why both overt C and null C are allowed in the contexts represented in (3) above, but null C is disallowed elsewhere. Many prior analyses account for null C distribution in purely syntactic or morphosyntactic terms:

- Stowell (1981) appeals to notions of Government and the Empty Category Principle (ECP) to unify licensing of the phonologically null complementizer with the licensing of other phonologically null syntactic elements.
- Pesetsky (1995) presents an account of null C licensing based on the idea that the null complementizer undergoes syntactic head-movement to adjoin to a local c-commanding lexical head.
- Similarly, Bošković & Lasnik (2003) propose that the null complementizer is a morphosyntactic affix which must attach to an immediately preceding lexical word.
- Bošković (2005) extends Bošković & Lasnik's (2003) analysis by characterizing the null complementizer as a syntactic enclitic, once again attaching to an immediately preceding lexical word.
- Lastly, Lohndal & Samuels (2013) develop a novel approach to the linearization of syntactic structures, along with limitations on the distribution of null syntactic terminals, in order to account for some aspects of null C licensing.

See An (2007a; b) for additional discussion of these and other accounts, such as Pesetsky & Torrego's (2001) feature-checking analysis.

Importantly, none of these approaches are able to fully capture the range of facts on null C distribution. They are all either forced to set some datapoints aside entirely or are obliged to introduce extra stipulations in order to cover errant data. In brief: Stowell's and Pesestky's accounts make incorrect predictions for null C licensing in the context of Right-Node-Raising. Bošković & Lasnik's null affix account runs into problems with Non-Bridge Verbs (see Bošković & Lasnik 2003: fn.1) and must also introduce a number of separate and distinct kinds of null complementizer to account for different configurations. Finally, Lohndal & Samuels' account cannot capture the facts for Clausal Subjects and Topicalized Clauses (see Lohndal & Samuels 2013: fn.6). Space does not permit a full overview of critiques of all prior theoretical accounts, but instead see Ormazabal (1995), Bošković & Lasnik (2003), and especially An (2007a; b) for relevant overviews and comprehensive evaluations.

For the purposes of this paper, I will choose to address the subsequent proposal by An (2007a; b), which outlines a distinct approach to null C distribution appealing to prosodic factors. An introduces the important insight that while the various constructions with *that* and Ø may have different syntactic derivations, *they must all be assigned prosodic structure*, and therefore prosody

offers a potential means of identifying a common through-line for the distribution of overt and null variants of C. An's account is based on a proposed principle of syntax-prosody mapping called the Intonational Phrase Edge Generalization:

(14) **Intonational Phrase Edge Generalization (IPEG):** The edge of an I-Phrase cannot be empty […] (An 2007a: 61).

This states that a phonologically null item cannot occupy the edge of an Intonational Phrase. Importantly, as discussed in the preceding section, An observes that many of the contexts in which null C is banned clearly involve obligatory parsing of the embedded clause into an IP, and therefore postulates that a principle like IPEG could ban alignment of null C at the left edge of this IP, which would create a misalignment between prosodic and syntactic constituents. An's account is attractive because it successfully covers many, if not all, of the facts from the preceding section. However, it nevertheless runs into some problems which make it undesirable as an account of null C distribution, as follows:

First, the IPEG is essentially a stipulation. It does not appear to follow from independent principles of grammar, aside from the notion that syntax and prosody should "match up". Lohndal & Samuels (2013: 5–6) present this criticism by asking what the motivation could be for a principle like the IPEG to exist. While IPs do generally line up with the edges of clauses (CPs), it is unclear why there should be an explicit requirement that they do so built into the grammar, especially in light of the fact that variation in IP-parsing is indeed observable, as with the cases of "integrated" IP-parsing in the previous section. See Lohndal & Samuels (2013) for additional discussion along these lines, as well as further critiques on conceptual grounds.

This leads to a second objection: the IPEG crucially requires strict isomorphy between syntactic and prosodic constituents at the level of the clause. This is a strong theoretical commitment which is not broadly accepted and does not clearly conform to our understanding of variable relations between syntax and prosody. Many models of syntax-prosody mapping assume some level of isomorphy in the initial syntax-to-prosody mapping (see, e.g., Match Theory, Selkirk 2009; 2011; Elfner 2012; 2015), but also allow for many adjustments to prosodic structure in response to external factors like discourse and information structure, length of constituent, rate of speech, and other elements of parsing and production (Nespor & Vogel 1986; Selkirk 1996). These adjustments yield many instances of non-isomorphy between syntactic and prosodic structure. Importantly, in order to maintain the IPEG, no such contextual adjustments are allowable. If null C is aligned to an IP-boundary in a resultant prosodic structure, the derivation is "stuck", without the possibility of, for example, selecting a different IP-parsing configuration to avoid or repair an aberrant null edge. The flexibility of prosodic phrasing is a longstanding observation in much prior work; see Selkirk (1978; 1984; 1996) and Nespor & Vogel (1986) for examples of variation in IP-parsing in English. As a result, we would expect that variable phrasing should

be available in many cases to resolve an itinerant null edge if IPEG is indeed a constraint on the mapping from syntax to prosody. In order for An's account to work, however, this option cannot be countenanced.

Taken altogether, I believe these issues speak against the legitimacy of IPEG, and so I will not adopt it here. However, as already noted, I will take as a starting point An's basic observations about the relevance of IP-boundaries and incorporate this into an account of null C distribution that still appeals to prosodic structure; namely, via the concept of prosodically conditioned allomorphy. This account, consisting of both a synchronic and diachronic component, is presented in the next section.

4 Prosodically conditioned allomorphy of C

The data on prosodic phrasing presented in Section 2 shows the following generalization, which I will call the IP-Boundary Generalization:

(15) **IP-Boundary Generalization:**

a. The contexts where null C is allowed do not show an obligatory IP-boundary before C, but instead an optional one. C appears *medially* within IP in such contexts: ($_{IP}$…C…).

b. The contexts where null C is disallowed show an obligatory IP-boundary before C. As a result, C always appears *initially* within IP in such contexts: ($_{IP}$ C…).

This basically encapsulates An's (2007a; b) observations about the correlation between the obligatory presence of an IP-boundary and prohibition on selection of null C. We have already seen, however, that An's account invoking syntax-prosody mapping and the IPEG is not ideal to account for this correlation.

I will now make an alternative proposal characterizing null C distribution as a case of prosodically conditioned allomorphy, an approach to the phenomenon which has not been explored thus far. The proposal consists of a synchronic component presenting how to understand the relation between the overt and null forms of the finite complementizer, followed by a diachronic component explaining the historical origin of the null allomorph, as well as the origins of the prosodic conditioning factors associated with it.

4.1 Synchronic proposal

I will start with the synchronic component:

(16) **Synchronic Proposal:** Both overt C (*that*) and null C (Ø) are allomorphs of the complementizer morpheme C which can be selected to represent C in the form of a sentence. Each allomorph is associated with a specific prosodic licensing context. Overt C may freely appear adjacent or non-adjacent to an IP-boundary, while null C may only appear medial to IP, non-adjacent to any boundary.

To formalize this, I will follow a line of work on phonologically-conditioned allomorphy which characterizes allomorph-selection as being governed by morphological subcategorization for a specific environment (Carstairs 1988; 1990; Paster 2006; Stanton 2021). Under these assumptions, each allomorph of C is associated with a subcategorization frame which defines the conditioning factors for that allomorph. Given the IP-Boundary Generalization above, the information instantiated in each subcategorization frame will be prosodic in nature, referring to prosodic boundaries, hence *prosodically* conditioned allomorphy.

The schemas below are formatted following Paster (2006) and Stanton (2021) and show the proposed conditioning factors for overt C and null C. Both frames also include information about the syntactic environment in which C must be inserted, namely as the head of a Complementizer Phrase (CP) combining locally with a Tense Phrase (TP) complement. Importantly, both schemas are also annotated with IP-boundaries, generally corresponding to the boundaries of CP:

(17) a. **Overt C (*that*)**
$(_{IP} \quad)$
$[\ (\ldots)\ \textit{that}_{C}\ [\ldots]_{TP}\ ]_{CP}$

b. **Null C (Ø)**
$(_{IP} \quad)$
$[\ \ldots\ \emptyset_{C}\ [\ldots]_{TP}\ ]_{CP}$

In (17a), the element (…) indicates that overt C is licensed to be inserted either immediately adjacent to a left IP-boundary, as in (4a) above, or non-adjacent to an IP-boundary, as in the integrated IP-parsing in (4b). This represents the optionality of selection of *that*. Rather than this being a simple case where one allomorph is licensed in a specific context and the other allomorph is licensed "elsewhere", selection of *that* is technically allowed "everywhere", while Ø is only allowed in one specific context—non-adjacent to a left IP-boundary (=17b)—a context which overlaps with the licensing context for *that*.

A reviewer raises the question of how to determine whether a prosodic boundary is obligatory or not and how that impacts selection of null C. A basic assumption here, which I believe is uncontroversial, is that certain sentence-types or constructions are "canonically" associated with specific prosodic forms, such as separate versus integrated IP-parsing, and that learners acquire this knowledge. So, as an example, Extraposition is canonically associated with a separate IP-parsing for the extraposed constituent. Once learners also acquire the prosodic conditioning for null C, they can then assess whether or not null C is appropriate for the canonical prosodic form of a given environment. As we have seen, many constructions are judged consistently by speakers to require a prosodic break before an embedded C, and therefore speakers are predicted to disprefer null C in those constructions, which they do. Crucially, this would be true even if, in a single specific utterance, a given construction was not pronounced with a clear prosodic break

before the embedded C. It is possible that the absence of a prosodic break would improve the acceptability of null C to some extent, but under the proposal here, the overall licensing of null C is determined by the licensing conditions that speakers have acquired in combination with the canonical prosodic form of constructions that they have also acquired.

As a final note, the distribution of overt C and null C should be understood as a case of prosodically conditioned *suppletive* allomorphy. This is because, according to typical standards by which allomorphic relations are established, the phonetic difference between overt and null forms of C are significant enough that they cannot be understood as a case of a purely phonological rule. Simply put, it is unlikely that there would be a lexically-specified rule for C stating "delete all segments" in order to derive Ø from *that* phonologically in the same way that allomorphs of the English plural *-s* suffix (*cat*[s]*, dog*[z]*, hous*[əz], etc.) are frequently assumed to be derived from some common morph via rules of Assimilation and Epenthesis. As a result, the synchronic relation between *that* and Ø should be understood as more similar to the relation between the verb *go* and its past tense form *went*, which are clearly not derived from a common underlying element in synchronic terms. Diachronically, however, there is a clear case to be made for linking *that* and Ø, bound up with the origins of the prosodic conditioning factors in the schemas above.

4.2 Diachronic proposal

A synchronic description and formalization of the allomorphy of the finite complementizer is all well and good, but how can it be justified as a useful and explanatory account? In order to accomplish this, we must explain the origins of the null allomorph itself, and we must also explain how allomorphs of C came to be associated with subcategorization frames which incorporate prosodic information. This leads to the diachronic component of the proposal, which has two parts.

(18) **Diachronic Proposal:**

a. Null C originates as the endpoint of visible processes of phonological weakening which affect the overt form *that* (/ðæt/) over historical time. The relevant processes are Vowel Reduction and Word Initial Deletion (/ðæt/ → /ðət/ → /ət/, /t/ → Ø).

b. The prosodic conditioning of overt and null C is a result of phonological weakening being allowed in prosodically "weak" environments (domain-medial position), while being prohibited in "strong" environments (domain-initial position). Over time, learners internalize correlations between weak vs. strong forms and their prosodic position, and these become encoded in the subcategorization frames of resulting allomorphs.

To start, we must explore the diachronic origin of null C in the history of English. In the Old English period (OE, 450–1066CE), the finite complementizer *þæt* had essentially the same syntactic function as its present-day form, but did not show a null counterpart, except in some limited "quotative" constructions with a form of the verb *say*, attested in late OE or early Middle

English (ME, 1150–1450CE). It is not until the late ME period that clear attestations of null C can be identified following matrix verbs and after nouns in restrictive relative clauses. These attestations are more likely to appear in texts which approximate speech, rather than formal register, such as sermons, comedies, and private letters (Rissanen 1991: 279). Frequency of null C continues to increase in Early Modern English (1450–1700CE), but drops in the 1800s, likely due to rising prescriptive attitudes (Rissanen 1991; Suárez-Gómez 2000). See Conde-Silvestre & Calle-Martín 2015 for further discussion.

Importantly, null C does not arise *ex nihilo* in late ME. In fact, it emerges alongside a number of variant forms, variously spelled *þæt, þat, þatte, þet, þt, at,* and *atte*. The last two forms are of particular interest since they provide overt evidence for processes of phonological reduction which are known to be at work in ME but which are not always expressed in spelling. These are general processes of phonological weakening which continue to be act work in Present-Day English (PDE): Vowel Reduction and Word Initial Deletion (or Lenition). I will briefly describe both of these as they manifest in PDE before returning to the historical situation.

Vowel Reduction is a phonological process which applies generally to unstressed vowels in English words, reducing them to schwa [ə]. It also optionally affects monosyllabic function words (prepositions, auxiliary verbs, determiners, pronouns, etc.), which do not have inherent word-stress, and is responsible for well-known strong/weak alternations in the form of function words, such that they may be pronounced with full vowels or with reduced vowels, as shown in (19) for *can, to, for* (Kaisse 1983; Selkirk 1996) and the complementizer *that* itself (Kandybowicz 2007; Tyler 2019).

(19) a. Sue **can** carve pumpkins. /kæn/ ~ /kən/
b. Sue drove **to** London. /tu/ ~ /tə/
c. Sue drove **for** hours. /foɹ/ ~ /fɻ/
d. I believe **that** Sue likes linguistics. /ðæt/ ~ /ðət/

Vowel Reduction in unstressed environments is attested in early Middle English, although it is not always reflected in spelling. In Present-Day English, the reduced form of *that* is sometimes written *th't*, as shown in (20):

(20) a. That is th' man, **th't** said that you *saw* him. (Bronson 1845: 49)
b. And more ov'r in opyn plae' said **th't** he would by ye said craft of enchauntement, wychecraft, and socerye, wyrke yo'r said suppliant his nekke to breke … (Polsue 1867: 86)
c. Ye heerd the las' witness say **th't** he sold me a six-oun' picker'l to-day … (Mott 1895: 63)

Although Vowel Reduction is an important step on the path toward phonological weakening, it is not necessarily a prosodically conditioned process of the kind we are interested in here.

The only relevant constraint on Vowel Reduction is when a function word occupies phrase-final position, with no structural complement following it. Prosodic accounts of this constraint have been proposed (see, e.g., Selkirk 1996); however, such accounts will not be relevant to the current proposal given that phrase-final position is never possible for C due to the unavailability of processes of structural displacement or ellipsis for the Tense Phrase complement of C in English. As a result, any potential prosodic constraint related to phrase-finality will not apply to C. As such, let's turn now to the second significant rule of phonological reduction, which affects segments in word-initial positions.

Word Initial Deletion is a phonological process which is more consistently restricted to function words.[2] Deletion applies to word-initial /h/, /w/, and /ð/ on functions words like auxiliary verbs, complementizers, and pronouns (Labov 1969; Zwicky 1970; Inkelas & Zec 1993; McElhinny 1993).

(21) a. Susan **has** visited today. /hæz/ ~ /æz/
b. Susan **would** rather travel. /wʊd/ ~ /ʊd/
c. Susan's trip was better **than** Morgan's. /ðɛn/ ~ /ɛn/

Vowel Reduction and Word Initial Deletion are distinct processes, although they frequently co-occur. See Labov (1969: 727) and Zwicky (1970: 334) for demonstrations of the independent application of these processes.

Of particular relevance here is the process of /ð/-deletion shown in (21c), which also notably applies to the complementizer *that*. Let's return now to the historical picture at this point. Deletion of /ð/ on earlier forms of C is also attested by late Middle English. (22) shows examples from the *Alphabet of Tales* (de Besançon 1905), which represents a northern late Middle English variety (c1450CE) and contains many examples of *þat* manifesting as *at* when it heads a clause following an embedding verb or a relativized noun:

(22) a. *And he **saw at** sho wolde not lett to ete for hym*
"And he **saw that** she would not stop eating for him" (de Besançon 1905: I.17)
b. *Thow **Lord at** made me, hafe mercie on me!*
"Thou **Lord that** made me, have mercy on me!" (de Besançon 1905: I.3)

The form *at* became systematically grammaticalized in northern varieties of Middle English and continues to appear in northern varieties of British English up until the present day. Wright (1903), for example, dedicates a full entry in the *English Dialect Dictionary* to complementizer-*at*,

[2] A similar process of Lenition applies in the same contexts as Word Initial Deletion, as well as word-internally in unstressed syllables, affecting aspirated stops like /k/ and /t/. Lenition yields deaspiration for /k/ and "flapping" for /t/, shown in (i) (Selkirk 1984; 1996; Cooper 1991; 1994).

(i) a. Susan **can** carve pumpkins. /kʰæn/ ~ /kæn/
b. Susan drove **to** London. /tʰu/ ~ /ɾu/

with additional examples from varieties of British English. (23a) shows another example from late Middle English (~1475CE), followed by more recent examples from Scottish English (=23b-c), along with a comparative example from Scots, which diverged in the early Middle English period (=23d). A form of complementizer-*at* also appears in varieties of English in North America, although it has not grammaticalized as such and is rarely signified in spelling. Two examples from late 1800s American English with the spelling *'at* are shown in (23e):

(23) a. …*quhilkis sall serche and se all* ***wirkis at*** *the craftismen wirkis.*
"which shall search and see all **works that** the craftsmen work." (Marwick 1869: I.31)

b. …*the* ***wumman ăt*** *ye ken hyr sun.*
"the **woman whose** son you know"
Literal: "the **woman that** you know her son" (Murray 1873: 196)

c. …an' I **saw 'at** he was waitin' for me (Barrie 1889/1898: 250)

d. He hed **sein at** it wesna a bad reddin up o the maitter (Forde 1996: 142)

e. 'Nd every one **thought 'at** he'd kill himself there— […] (King 1898/1906: 179)
He **says 'at** tomorrer […] He has a notion to / Try it again

To round out the picture, I also provide the following judgements from my own dialect of Western American English, which have been informally confirmed by consultation with other speakers of varieties of both American and British English. According to these judgements, /ð/-deletion continues to be available, and combined with Vowel Reduction, yields a set of forms ranging from /ðæt/ to /ðət/ to /ət/ and even /t/, depending on rate-of-speech and surrounding segmental context. I also add Ø to the list of options in the list below, completing the full set of phonological forms that C can take:

(24) I believe { **/ðæt/** / **/ðət/** / **/(ə)t/** / **/Ø/** } Sue likes linguistics.

The goal here has been to demonstrate that the emergence of null C in the history of English was not an isolated event, but instead corresponded with the emergence of a range of reduced forms of overt C, many of which are attested in the historical record and are still attested in PDE. These forms emerge as a result of phonological weakening processes like Vowel Reduction and Word Initial Deletion, and in some cases the dual application of these processes can result in heavy reduction of C. It is a small step from this observation to the first part of the Diachronic Proposal in (18): that null C is the end result of these reduction processes, i.e. reduction to zero.

Let's turn now to the second part of the Diachronic Proposal, which addresses the origins of the prosodic conditioning that is incorporated into the proposed schemas for allomorphy selection

in Section 4.1. There is an established correlation between the application of phonological weakening processes and the prosodic environment of a word/segment. In particular, domain-medial positions within prosodic constituents (i.e. non-adjacent to any prosodic boundary) are known to be "weak" and associated with phonological weakening, while domain-initial positions are "strong", disallowing weakening. These correlations are noted as early as Lass & Anderson (1975) and Escure (1977) (see also Cooper 1991; 1994; Pierrehumbert & Talkin 1992; Fougeron & Keating 1997; Beckman 1998; Cho & Keating 2001; Fougeron 2001; Keating et al. 2004; White & Turk 2010).

To illustrate more concretely: Selkirk (1996) and Ito & Mester (2009a; b) observe that segmental weakening in the form of Lenition of voiceless stops like /t/ and /k/—a process in the same category as Word Initial Deletion (see footnote 2)—is constrained at the initial boundaries of most prosodic constituents. In (25a) below, we see that Lenition is allowed for the /t/-segment associated with the preposition *to*, yielding [ɾ] ("flapping"), but not for the initial /t/-segment of the lexical noun *town*, which is parsed as a Prosodic Word (PW). In (25b-c), /t/ is aligned to the start of an IP and to the start of an Utterance (U), respectively, and Lenition is prohibited in both cases.

(25) a. Sue drove **t**o [ɾə] **t**own [**t**^{h}a͡ʊn]. … **t**o ($_{PW}$ **t**own)
b. Sue drove, she said, **t**o [**t**^{h}ə] town. … she said ($_{IP}$ **t**o town)
c. Where did she drive? – **T**o [**t**^{h}ə] town. ($_{U}$ **T**o town)

Crucially, the same prosodic restrictions apply to Word Initial Deletion, illustrated by /h/-deletion, which is allowed on the auxiliary verb *has* unless it is aligned to an initial prosodic boundary (Labov 1969; Zwicky 1970):

(26) a. Sue has [æz] helped [**h**ɛlpt] me today. … **h**as ($_{PW}$ **h**elped)
b. Sue, along with Max, **h**as **[h**æz] arrived. … along with Max ($_{IP}$ **h**as arrived)
c. Has Sue called today? – **H**as [**h**æz**]** she? ($_{U}$ **H**as she?)

Constraints on processes of weakening like Lenition and Word Initial Deletion are the outcome of the well-established crosslinguistic phenomenon of Domain-Initial (Articulatory) Strengthening (DIS), whereby the articulation of consonantal segments is stronger at initial prosodic boundaries and weaker in medial positions. DIS has been studied via experimental work involving electropalatography, measuring the relative degree of linguopalatal contact when speakers pronounce segments in different prosodic positions (Fougeron & Keating 1997; Cho & Keating 2001; Fougeron 2001; Keating et al. 2004; see also Keating 2003 and Cho 2011 for overviews). The effects of DIS have been identified across a wide range of languages, such as Estonian (Gordon 1999), Korean (Cho & Keating 2001), French, Taiwanese (Keating et al 2004; also Tabain 2003; Tabain et al. 2003), and Japanese (Onaka 2003). The following two findings are relevant (adapted from Keating 2003):

(27) **Domain-Initial Strengthening**

a. Articulatory strength correlates with adjacency to the leading edge of a prosodic constituent. Segments in domain-initial position (x …) are articulated more strongly than segments in domain-medial position (… x …).

b. Articulatory strength correlates with the level of prosodic constituent in the "Prosodic Hierarchy" (Selkirk 1972; Nespor & Vogel 1986). The higher the level, the higher the likelihood for stronger articulation. A segment aligned to the edge of an Utterance is stronger than a segment aligned to the edge of an Intonational Phrase, which is stronger than a segment aligned to the edge of a Phonological Phrase, a Prosodic Word, etc.

For our purposes here, we can set up a contrast between Domain-Initial Strengthening, which creates a resistance to the emergence of weakened forms in domain-*initial* positions, and Domain-Internal Weakening (DIW) which allows for the emergence of weakened forms in domain-*internal* or domain-*medial* positions, i.e. in any position that is not aligned to an initial boundary. The second component of the Diachronic Proposal simply states that, over time, speakers of English observe correlations between the prosodic position of function words and their strong/weak forms. Strong forms are associated with domain-initial prosodic positions, while weak forms (especially those subject to segmental processes like Lenition and Word Initial Deletion) are associated with domain-internal prosodic positions.

In the case of the complementizer *that*, as DIW causes the complementizer to undergo phonological reduction (eventually to zero in the late Middle English period) prosodic position becomes internalized as "prosodic conditioning" for this form, and the zero form emerges as a distinct allomorph correlating with medial positions. As such, the proposal here is that the resulting distribution of null C is the outcome of the kind of process noted by Labov (1969: 755–757), whereby instances of rule-governed phonological reduction eventually lead speakers to postulate true underlying zero forms, as with, for example, zero forms of the verbal auxiliaries *have* and *been* in African American English and other English varieties. Once acquired, the null allomorph is still conditioned by the environments in which learners were exposed to it, although it could technically be selected optionally in other environments, possibly with degraded or dispreferred outcomes.

Note that null C is distinct from phonologically reduced forms like /ðət/ and /ət/, which can be derived systematically from the full form /ðæt/ by general rules of phonological reduction like Vowel Reduction and Word Initial Deletion and need not be characterized by overt prosodic conditioning in a schema. In contrast, null C is a separate allomorph of C—not simply a phonological reduction—since there is no systematic rule of phonological weakening or deletion in English that can fully derive Ø from, e.g., /ðæt/. In other words, null C is not being reduced to zero every time that it appears in speech. It originates diachronically as a reduction, but synchronically, it exists as a null allomorph of C with the conditioning environments preserving the prosodic environments in which phonological reduction has historically occurred.

4.3 Summary

Table 1 summarizes the account by showing the relevant historical stages of development for the complementizer and which forms were available at which times, locating the emergence of null C at some time in the Middle English period.

Historical Stage	Full	Reduced	Zero
Old English	þæt		
Early Middle English	þæt, þat, þatte, þet,	þt, at, atte	
Late Middle English	þæt, þat, þatte, þet,	þt, at, atte	Ø
Present-Day English	that	th't, 'at, 't	Ø

Table 1: Development of allomorphs of C in English.

The arrow indicates the diachronic relation between the reduced forms of the complementizer in Early Middle English, which then (according to the Diachronic Proposal) provided the stepping-stone toward the emergence of a zero form by Late Middle English. From then onward, according to the Synchronic Proposal, this zero form would be constrained or allowed by the same prosodic conditioning factors of DIS and DIW that historically constrained or allowed the reduced forms. In synchronic terms, prosodically conditioned allomorphy is represented by prosodic information incorporated into allomorphic schemas for overt C and null C; namely, the presence vs. absence of a preceding prosodic boundary.

5 Discussion

In this section, I will discuss some consequences of the synchronic and diachronic proposals and also address specific aspects of the data in Section 2 and how it justifies a prosodic approach to null C distribution.

First, note that the synchronic account as presented focuses on the level of IP only. However, DIS has been shown to occur in domain-initial position at *all* prosodic levels—not just at IP—and is stronger at "higher" levels, see (27b) above. This is relevant to models of syntax-prosody because it leads to predictions about where prosodic boundaries should be postulated. For example, in cases where integrated IP-parsing occurs, illustrated in (4b) (repeated below), we

know that both Word Initial Deletion and null C are both available in this context, and therefore we must additionally assume that C is not only medial to IP (non-adjacent to an IP-boundary) in (4b) but also medial to lower-level prosodic constituents, such as a Phonological Phrase (PhP).

(4) a. ($_{IP}$ Our next door neighbor truly believes) ($_{IP}$ that black cats bring bad luck)
b. ($_{IP}$ Our next door neighbor truly believes that black cats bring bad luck)

Although I will not develop a full account here, there are generally two options for the prosodification of function words at levels below IP: (i) C might be integrated into a lower prosodic constituent on its left, having the status of a prosodic *enclitic*, as in (28a) below, or (ii) C may be integrated into a lower prosodic constituent on its right, having the status of a prosodic *proclitic*, as in (28b). In the latter, case, C would be aligned to the initial prosodic boundary of a PhP, and therefore both Word Initial Deletion and null C itself are predicted to be disallowed. This cannot be the case, however, if we understand the availability of deletion and null C to specifically correlate with medial position and therefore the absence of a preceding prosodic boundary. Thus, for the account of null C presented here, I must assume that lower-level prosodic organization of C in sentences with integrated IP-parsing looks something like (28a), rather than (28b). Both of these sentences show PhP-parsing for English, which is generally known to at least parse subject Noun Phrases and Verb Phrases into separate PhPs (Nespor & Vogel 1986).

(28) a. **Enclitic C**
($_{PhP}$ Our next door neighbor) ($_{PhP}$ truly believes **that**) ($_{PhP}$ black cats) ($_{PhP}$ bring bad luck)
b. **Proclitic C**
($_{PhP}$ Our next door neighbor) ($_{PhP}$ truly believes) ($_{PhP}$ **that** black cats) ($_{PhP}$ bring bad luck)

Notably, the concept of enclitic complementizers has precedent in the syntactic and prosodic literature; see, for example, Schütze (1994), Pesetsky (1995), Bošković & Lasnik (2003), and Bošković (2005). Each of these works posit that when an embedded clause is complement to a lexical word (*believes* in (28)), the embedded C bears a closer structural relation to that lexical word than to the embedded clause itself. Schütze (1994: 463–464) suggests on the basis of data from Serbo-Croatian that an embedded C can be prosodically organized into the same IP as the selecting matrix verb it is adjacent to. Pesetsky (1995), Bošković & Lasnik (2003) and Bošković (2005) make these ideas more explicit for English data by proposing that null C in particular is either a suffix or a syntactic enclitic, attaching to a lexical category to its left and, in the case of Bošković (2005), subject to various restrictions imposed by intervening prosodic boundaries. More broadly, a number of papers have explored evidence for more widespread enclisis in the prosodification of English function words, including Wheeldon & Lahiri (1997), Lahiri & Plank (2010), Ito & Mester (2019), and Tyler (2019). The reader is directed to these works for further information. For our purposes here, the important thing to reiterate is that an account which appeals to the differential effects of DIS and DIW must at least allow the option of enclisis for

C in order to maintain correct predictions for the occurrence of Word Initial Deletion and the prosodic conditioning of null C.

I will now move on to address some specific aspects of the data in Section 2 which highlight the advantages of a prosody-based account such as the one presented here. First, perhaps the most intuitive aspect of the account is the way that it captures null C licensing with Clausal Subjects and Topicalized Clauses, repeated below, both of which position C in sentence-initial position.

(5) a. **{That, *Ø}** the teacher was lying was hardly obvious. Clausal Subject
b. **{That, *Ø}** John likes Mary, Jane didn't believe. Topicalized Clause

It has already been mentioned that selection of null C in these constructions is strongly unacceptable. In fact, of the many constructions discussed so far, selection of null C in these sentences may be the most consistently cited example of null C prohibition in the literature, and I am unaware of any example where null C is actually attested in such an environment, including in corpus work by Kearns (2007) which does otherwise show variability for many other constructions. This is, in fact, predicted by the account of this paper, given the effects of DIS. Since in most prosodic frameworks, the beginning of a sentence will be parsed as an IP-boundary at least, and possibly as an Utterance-boundary as well (a higher prosodic constituent than IP), sentence-initial position is predicted to be the strongest prosodic position possible, and therefore the most resistant to phonological reduction. Consequently, this position would also be expected to yield the highest level of unacceptability should the prosodically conditioned null allomorph be inserted there.

This account has substantial advantages compared to other proposals, some of which were already briefly noted in Section 3. For example, proposals by Lohndal & Samuels (2013), which attempt to derive null C licensing via linearization mappings from syntax, simply leave clause-initial null C aside as an unexplained phenomenon within their account; almost certainly a mistake given the consistent prosodic patterns pointed out here. In contrast, Bošković & Lasnik (2003) propose a morphosyntactic account characterizing null C as an affix which must attach to an adjacent lexical word. This successfully derives many of the adjacency effects on null C licensing seen in the data, including prohibition of null C in clause-initial position, where there is no preceding lexical word for null C to affix to. However, the authors do note that their account provides no explanation for why affixal null C cannot simply attach to an adjacent nominal subject immediately to the right of C, e.g., *teacher* or *John* in (5) above (Bošković & Lasnik 2003: fn.31). While not fatal to their account, it does require more bespoke complexity in the subcategorization requirements of the supposed null C affix and again fails in large part to identify the common pattern tying clause-initial null C to other cases of null C prohibition.

Next, let's address null C prohibition under Extraposition and Parenthetical insertion as in (6), repeated below.

(6) a. I believe very strongly **{that, *Ø}** Sue likes linguistics. Extraposition
b. I believe, based on intuition, **{that, *Ø}** Sue likes linguistics. Parenthetical

It should be noted that Extraposition is particularly difficult to capture in prior accounts that appeal to the adjacency relation between a lexical word and embedded C (e.g., Bošković & Lasnik 2003; Lohndal & Samuels 2013) because even under Extraposition, embedded C can remain adjacent to some lexical word, such as an adverb (as with *strongly* in (6a)) or the NP complement of a preposition (*I said to **Kay that** Sue likes linguistics*). Accounts which tie the licensing of null C to some aspect of underlying syntax, such as Stowell's (1981) appeal to Proper Government or Pesetsky's (1995) appeal to syntactic head-raising, actually fare better in this respect because a clear connection can be drawn between the syntactic rule of Extraposition (however it is characterized) and subsequent licensing or non-licensing of null C. At the same time, a prosodic account is also able to draw clear connections, given the prosodic effects of Extraposition, which prevent IP-integration.

Turning now to Parenthetical insertion, it should be noted that parentheticals are rarely addressed in prior accounts of null C distribution, but may, in fact, be the strongest argument in favor of a prosodic approach. This is because one of the defining criteria for parentheticals is that they may be freely inserted at a number of points in the structure of a sentence and do not "disrupt" the structure around them. Parentheticals are generally assumed to attach externally to the unit that hosts them, either via adjunction or some other operation (see Emonds 1973; 1976; 1979; McCawley 1982; Haegeman 1991; Potts 2004 for a variety of approaches to parenthetical attachment). This is significant because, if null C is allowed in one sentence but disallowed in the same sentence with a parenthetical inserted, this suggests that the licensing factor for null C in these sentences is not strictly syntactic or semantic—not strictly a part of the "underlying" form of a sentence. In other words, we cannot generally assume that a parenthetical disrupts or destroys an underlying structural configuration, because parentheticals are understood to be external to such configurations. At the same time, it is well known that parentheticals do indeed disrupt the *prosodic* structure of a sentence, as discussed above, breaking up an otherwise-integrated IP structure. If parentheticals affect prosodic structure, but not syntactic or semantic structure, and null C is prohibited when a parenthetical is inserted (as in (6b)), this strongly suggests that the licensing of null C should be located in the prosodic form of a sentence, rather than elsewhere.

Looking at the whole of the data presented in Section 2, it should also be emphasized that the constructions which disallow null C are heterogenous in nature. They include constructions where an embedded clause occupies a special syntactic position (Clausal Subjects) or a "derived" position related to discourse/pragmatic phenomena (Topicalized Clauses, Extraposition), the

insertion of other phrases in relation to the embedded clause (Parentheticals), different underlying argument structures and structural relations between the embedded clause and other categories (Non-Bridge Verbs, Noun Complement Clauses), the application of rules of coordination and ellipsis, and the assignment of focus (Right-Node-Raising, Gapping, Wh-clefts). Looking at the problem from a high altitude, however, it is worth asking what linguistic system could tie together all of these phenomena and end up yielding the same effect (prohibition of null C). This highlights the explanatory power of the prosodic account here: despite their diverse syntactic, semantic, and discourse/pragmatic properties, all of the relevant environments must be assigned prosodic structure, and it is this shared structure which determines null C distribution through the medium of prosodically conditioned allomorphy.

Having established some reasons to believe the prosodic account is preferable, I believe that the proposals of this paper can also serve as the starting point for a broader program of research on the boundary between prosody, syntax, and morphology, since it sets up a series of testable predictions for future work to address. First, note that null C is only one manifestation of the larger phenomenon of null allomorphy in English. The account in this paper predicts that null allomorphy may (and even should) also emerge in other prosodic environments. This may allow for extensions of the account to reduced and null counterparts of other function words in varieties of English, such as reductions of verbal auxiliaries. In the sentences below, both reduced and null forms of the auxiliaries are licensed except when a prosodic boundary occurs before the auxiliary (illustrated via Parenthetical insertion, with comma punctuation indicating prosodic breaks):[3]

(29) a. When **{are, -‘re, Ø}** you coming back?
b. When, pray tell, **{are,** ***-‘re,** ***Ø}** you coming back?

(30) a. What **{did, -‘d, ?Ø}** she say?
b. What, if you don’t mind me asking, **{did,** ***-‘d,** ***Ø}** she say?

(31) a. What **{have, -‘ve, Ø}** you been doing?
b. What, perish the thought, **{have,** ***-‘ve,** ***Ø}** you been doing?

A reviewer asks what might constrain rampant reduction to zero in the case of other function words, if it is indeed available and conditioned by prosodic context. To answer this, note that reduction to zero has practical consequences due to loss of information from the surface form. It is possible that null C is licensed over time due to the “redundancy” of C for marking the clause-edge in specific contexts. When a clause is embedded under high-frequency clause-taking verbs

[3] These cases of reduced and null allomorphy must be distinguished from clause-initial ellipsis of pronominal subjects and auxiliaries (Are you *coming back?* Did *she really say that?* Have you *gone to the store lately?*), which is a distinct phenomenon in English (Napoli 1982; Zwicky & Pullum 1983).

like *say* or *think,* the existence of the embedded clause becomes highly predictable, and this readily allows reduction to zero over time. The same may be true for verbal auxiliaries *have* and *be* referenced above, whose presence is also signaled by participial forms of the verbs they select for. Other function words like the prepositions *to* and *for* do not have this functional redundancy. They convey semantic information that cannot be recovered purely via the predictability of the context around them, and so they show phonological reduction, but have so far resisted reduction to zero.

An additional important aspect of any prosodic account is that prosodic phrasing in general (and IP-parsing in particular) is variable, subject to a number of factors like the length of a phrase or clause, the prosodic prominence of elements, speech style, and rate of speech (Nespor & Vogel 1986: 193). This creates predictions for the manifestation of null allomorphs in different environments. For example, a longer or more complex embedded clause is more likely to induce IP-boundaries compared to a shorter clause, and this should potentially have consequences for the selection of a null complementizer: the shorter the clause, the higher the likelihood of prosodic integration, and therefore of null C licensing). Consider also the fact that focus prominence (shown by SMALL CAPS below) has been observed to induce a prosodic juncture after a focused word (Selkirk 2002):[4]

(32) Sue THOUGHT that Mary would win, and she DID.
→ (Sue THOUGHT) (that Mary would win) (and she DID)

This predicts that null C should be dispreferred after a focused word. My judgements on this are not entirely clear. Null C seems to be available, but I do not judge the other reduced forms of the complementizer to be as acceptable. At the same time, this is complicated by the nature of focus-assignment, which also implies that post-focal material is presuppositional and/or "given". Interactions of givenness with null C licensing would also need to be explored in a future study, along with the impact of downstep and post-focal pitch compression for the determination of prosodic boundaries. Lastly, the nature of the prosodification of function words via proclisis or enclisis (as discussed at the start of this section) will also have to be more firmly established.

To conclude this section, any account of the English finite complementizer should mention the classic puzzles of the *that*-trace effect (obligatory null C with long-distance subject extraction) and the so-called anti-*that*-trace effect (obligatory overt C with short-distance subject extraction).

(33) a. Who did Sue think ***that/Ø*** left? *that*-trace effect
b. Sue bought the book **that**/***Ø** was written by Kurt. anti-*that*-trace effect

In the case of the *that*-trace effect, the account in this paper predicts that null C should be available at least, but the prosodic account does not provide a mechanism that would *enforce*

[4] Thanks to an anonymous reviewer for bringing this prediction to my attention.

(rather than simply allow) the selection of null C. The anti-*that*-trace effect is more noteworthy, since it appears to prohibit null C in a context where prosodic integration is otherwise available (non-restrictive relative clauses). It is notable, however, that reduced forms of C *can* be selected in (33b), as expected in a prosodically weak environment. The prohibition is simply that C cannot be fully null. Given the application of short-distance subject extraction in these sentences, I believe other external factors (antilocality, avoidance of parsing ambiguity, etc.) may explain this peculiarity (see Bickerton 2014 and Douglass 2017 for proposals).

6 Conclusion

In this paper, I have argued for an account of null C distribution based on the observation that the contexts which prohibit selection of null C also show an obligatory IP-boundary preceding the position of the complementizer, while contexts which allow null C do not show an obligatory boundary. This is formalized in a synchronic proposal that null C is a prosodically conditioned allomorph of the finite complementizer, licensed to appear in domain-medial position within a prosodic constituent. Overt C, on the other hand, is licensed to appear in either domain-initial or domain-medial position. This account is complemented by an additional diachronic proposal explaining how such a null allomorph gradually arose in the history of English through phonological reductions of the full form *that*, and how this particular allomorph ultimately came to be associated with specific prosodic environments as a result of the principles of Domain-Initial Strengthening and Domain-Internal Weakening, which govern the distribution of phonological reduction within prosodic constituents.

This account builds upon a line of work from Pesetsky (1995) through Bošković (2001), Bošković & Lasnik (2003), and An (2007a; b), emphasizing the relevance of prosody for explaining the distribution of nul C, rather than, e.g., purely syntactic properties of sentences. Each of these works identify important factors that have led to the core proposal here, from adjacency between the complementizer and other lexical words to the relevance of prosodic patterns which ultimately tie together the different constructions which disallow null C. This paper takes the final step to argue that the entirety of null C distribution can be captured by prosodic factors, once understood both diachronically and synchronically in the context of phonological weakening processes, principles of intonational phrasing in English, and the possibility of prosodically conditioned allomorphy.

Acknowledgements

I am grateful to Elly van Gelderen, Andrew Carnie, and Kathryn Pruitt for valuable feedback and comments on the earliest versions of this paper, which originated as an outgrowth of my dissertation. I would also like to thank the participants of the ASU Syntax Reading Group, as well as attendees at the 2020 LSA Annual Meeting for comments on subsequent drafts and a poster-version of the paper. Finally, I am grateful to four reviewers for the commentary and criticism they provided, which helped to substantially sharpen the argumentation and organization of the paper. Any remaining errors are my own.

Competing interests

The author has no competing interests to declare.

References

An, Duk-Ho. 2007a. *Syntax at the PF interface: Prosodic mapping, linear order, and deletion.* Storrs, CT: University of Connecticut dissertation.

An, Duk-Ho. 2007b. Clauses in noncanonical positions at the syntax-phonology interface. *Syntax* 10. 38-39. DOI: https://doi.org/10.1111/j.1467-9612.2007.00096.x

Aoun, Joseph & Hornstein, Norbert & Lightfoot, David & Weinberg, Amy. 1987. Two types of locality. *Linguistic Inquiry* 18. 537–577.

Astruc, Lluisa. 2005. *The intonation of extra-sentential elements in Catalan and English*. Cambridge: University of Cambridge dissertation.

Barrie, James M. 1898. *A Window in Thrums*. New York: Dodd, Mead and Company. (Original work published in 1889).

Beckman, Jill. 1998. *Positional faithfulness.* Amherst, MA: University of Massachusetts dissertation.

Bickerton, Derek. 2014. *More than Nature Needs.* Harvard: Harvard University Press. DOI: https://doi.org/10.4159/9780674728523

Bošković, Željko. 2001. *On the Nature of the Syntax-Phonology Interface: Cliticization and Related Phenomena*. Amsterdam: Elsevier. DOI: https://doi.org/10.1163/9780585474250

Bošković, Željko. 2005. *Null C in English as an enclitic.* University of Connecticut unpublished manuscript. URL: https://boskovic.linguistics.uconn.edu/wp-content/uploads/sites/2801/2019/05/nullCridjanovic6.pdf

Bošković, Željko & Lasnik, Howard. 2003. On the distribution of null complementizers. *Linguistic Inquiry* 32(4). 527–546. DOI: https://doi.org/10.1162/002438903322520142

Bronson, Cotesworth P. 1845. *Elocution, or, Mental and Vocal Philosophy*. Louisville, KY: Morton & Griswold.

Carstairs, Andrew. 1988. Some implications of phonologically conditioned suppletion. *Yearbook of Morphology* 1988. 67–94. DOI: https://doi.org/10.1515/9783112329528-006

Carstairs, Andrew. 1990. Phonologically conditioned suppletion. In Dressler, Wolfgang U. & Luschützky, Hans C. & Pfeiffer, Oskar E. & Rennison, John R. (eds.), *Contemporary Morphology*, 17–23. New York: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110874082.17

Cho, Taehong & Keating, Patricia. 2001. Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics* 29. 155–190. DOI: https://doi.org/10.1006/jpho.2001.0131

Collins, Peter C. 1991. *Cleft and pseudo-cleft constructions in English.* London & New York: Routledge.

Collins, Peter C. 2006. *It*-clefts and *wh*-clefts: Prosody and pragmatics. *Journal of Pragmatics* 38. 1706–1720. DOI: https://doi.org/10.1016/j.pragma.2005.03.015

Conde-Silvestre, J. Camilo & Calle-Martín, Javier. 2015. Zero *that*-clauses in the history of English. A historical sociolinguistic approach (1424–1681). *Journal of Historical Sociolinguistics* 1(1). 57–86. DOI: https://doi.org/10.1515/jhsl-2015-0003

Cooper, Alexander. 1991. *An Articulatory Account of Aspiration in English.* New Haven, CT: Yale University dissertation.

Cooper, Alexander. 1994. *Aspiration in English: How should it be defined and where should it be described?* Ann Arbor: University of Michigan manuscript.

de Besançon, Etienne. 1905. *Alphabet of Tales: An English 15th Century Translation of the Alphabetum Narrationum of Etienne de Besançon, from additional MS. Add. 25719 of the British Museum* (M. M. Banks, ed.). London: Kegan Paul, Trench, Trübner, & Co., Ltd.

de Cuba, Carlos & Urogdi, Barbara. 2010. Clearing up the 'facts' on complementation. *University of Pennsylvania Working Papers in Linguistics* 16(1). 41–50.

Declerck, Renaat. 1984. The pragmatics of *it*-clefts and *wh*-clefts. *Lingua* 64. 251–289. DOI: https://doi.org/10.1016/0024-3841(84)90065-2

Declerck, Renaat. 1988. *Studies on Copular Sentences, Clefts and Pseudo-Clefts*. Leuven: Leuven University Press/Foris Publications. DOI: https://doi.org/10.1515/9783110869330

Dehé, Nicole. 2009. Clausal parentheticals, intonational phrasing, and prosodic theory. *Journal of Linguistics* 45. 569–615. DOI: https://doi.org/10.1017/S002222670999003X

Dehé, Nicole & Kavalova, Yordanka. 2007. Parentheticals: An introduction. In Dehé, Nicole & Kavalova, Yordanka (eds.), *Parentheticals*, 1–22. Amsterdam/Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/la.106.03deh

Den Dikken, Marcel. 2006. Specificational copular sentences and pseudoclefts. In Everaert, Martin B. H. & van Riemsdijk, Henk (eds.), *The Blackwell companion to syntax, vol. 4,* 292–409. Oxford: Blackwell Publishers. (Original manuscript circulated in 2001). DOI: https://doi.org/10.1002/9780470996591.ch61

Den Dikken, Marcel & Meinunger, André & Wilder, Chris. 2000. Pseudoclefts and ellipsis. *Studia Linguistica* 54(1). 41–89. DOI: https://doi.org/10.1111/1467-9582.00050

Douglass, Jamie. 2017. Unifying the *that*-trace and anti-*that*-trace effects. *Glossa: a journal of general linguistics* 2(1):60. 1–28, DOI: https://doi.org/10.5334/gjgl.312

Downing, Bruce. T. 1970. *Syntactic structure and phonological phrasing in English.* Austin, TX: University of Texas dissertation.

Downing, Laura. 2011. The prosody of 'dislocation' in selected Bantu languages. *Lingua* 121. 772–786. DOI: https://doi.org/10.1016/j.lingua.2010.11.006

Elfner, Emily. 2012. *Syntax-prosody interactions in Irish.* Amherst, MA: University of Massachusetts dissertation.

Elfner, Emily. 2015. Recursion in prosodic phrasing: evidence from Connemara Irish. *Natural Language & Linguistic Theory* 33(4). 1169–1208. DOI: https://doi.org/10.1007/s11049-014-9281-5

Emonds, Joseph E. 1970. *Root and structure-preserving transformations*. Bloomington, Indiana: Indiana University Linguistics Club.

Emonds, Joseph E. 1973. Parenthetical clauses. In Corum, Claudia & Smith-Stark, Cedrik & Weiser, Ann (eds.), *You take the high node and I'll take the low node*, 333–347. Chicago: Chicago Linguistic Society.

Emonds, Joseph E. 1976. *A transformational approach to English syntax: Root, structure preserving, and local transformations.* New York, San Francisco & London: Academic Press.

Emonds, Joseph E. 1979. Appositive relatives have no properties. *Linguistic Inquiry* 10(2). 211–243.

Erteschik-Shir, Nomi. 1973. *On the nature of island constraints.* Cambridge, MA: MIT dissertation.

Escure, Geneviève. 1977. Hierarchies and phonological weakening. *Lingua* 43. 55–64. DOI: https://doi.org/10.1016/0024-3841(77)90048-1

Feldhausen, Ingo. 2010. *Sentential form and prosodic structure of Catalan*. Amsterdam/Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/la.168

Féry, Caroline & Hartmann, Katherina. 2005. The focus and prosodic structure of German right node raising and gapping. *Linguistic Review* 22. 69–116. DOI: https://doi.org/10.1515/tlir.2005.22.1.69

Forde, Iain. W. D. 1996. *Paix Machine*. UK: Fons Scotiae.

Fougeron, Cécile. 2001. Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics* 29. 109–136. DOI: https://doi.org/10.1006/jpho.2000.0114

Fougeron, Cécile & Keating, Patricia. 1997. Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 106(6). 3728–3740. DOI: https://doi.org/10.1121/1.418332

Franks, Steven. 2000. A PF-Insertion Analysis of *that*. *Syntaxis* 3. 1–27.

Franks, Steven. 2005. What is *that*? *IU Working Papers in Linguistics* 5. 33–62.

Frascarelli, Mara & Hinterhölzl, Roland. 2007. Types of topics in German and Italian. In Schwabe, Kerstin & Winkler, Susanne (eds.), *On Information Structure, Meaning and Form*, 87–116. Amsterdam/Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/la.100.07fra

Frascarelli, Mara & Ramaglia, Francesca. 2009. *(Pseudo)cleft constructions at the interfaces.* Unpublished manuscript. URL: https://ling.auf.net/lingbuzz/000841

Gordon, Matt. 1999. The effect of stress and prosodic phrasing on duration, acoustic amplitude and air flow of nasals in Estonian. *UCLA Working Papers in Phonetics* 92. 151–159.

Guimarães, Max. 1998. PF reflexes of phase boundaries. *Contribution to Minimalist Issues Chat Room*, University of Maryland.

Gussenhoven, Carlos. 2004. *The Phonology of Tone and Intonation.* Cambridge Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511616983

Ha, Seungwan. 2008. *Ellipsis, right node raising and across the board constructions.* Boston University dissertation. DOI: https://doi.org/10.3765/bls.v34i1.3562

Haegeman, Liliane. 1991. Parenthetical adverbials: The radical orphanage approach. In Chiba, Shuki & Ogawa, Akira & Fuiwara, Yasuaki & Yamada, Norio & Koma, Osamu & Yagi, Takao (eds.), *Aspects of modern linguistics: Papers presented to Masatomo Ukaji on his 60th birthday*, 232–254. Tokyo: Kaitakushi.

Hale, Kenneth, and Elisabeth Selkirk. 1987. Government and tonal phrasing in Papago. *Phonology Yearbook* 4. 151–183. DOI: https://doi.org/10.1017/S0952675700000804

Hamlaoui, Fatima & Szendrői, Kriszta. 2017. The syntax-phonology mapping of intonational phrases in complex sentences: A flexible approach. *Glossa: A Journal of General Linguistics* 2(1):55. 1–31. DOI: https://doi.org/10.5334/gjgl.215

Hartmann, Jutta. M. 2006. Well, there's the list reading. In Dotlacil, Jakub & Gehrke, Berit (eds.), *UIL OTS Working Papers in Linguistics*, 1–15. Utrecht Institute of Linguistics.

Hartmann, Katherina. 2000. *Right Node Raising and Gapping: Interface Conditions on Prosodic Deletion.* Amsterdam/Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/z.106

Herment, Sophie & Leonarduzzi, Laetitia. 2012. The pragmatic functions of prosody in English cleft sentences. In Ma, Qiuwu & Ding, Hongwei & Hirst, Daniel (eds.), *Proceedings of Speech Prosody 2012*, 713–716. Shanghai: Tongji University Press.

Higgins, Francis R. 1979. *The Pseudo Cleft Construction in English*. New York: Garland.

Hornstein, Norbert & Lightfoot, David. 1991. On the nature of lexical government. In Freidin, Robert (ed.), *Principles and parameters in comparative grammar*, 365–391. Cambridge, MA: MIT Press.

Inkelas, Sharon. 2014. *The interplay of morphology and phonology*. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199280476.001.0001

Inkelas, Sharon & Zec, Draga. 1993. Auxiliary Reduction without Empty Categories: A Prosodic Account. *Working Papers of the Cornell Phonetics Laboratory* 8. 205–253.

Ishihara, Shinichiro. 2014. On Match constraints. Poster presented at the *Workshop on the prosodic hierarchy in a typological perspective*, Stockholm University.

Ito, Junko & Mester, Armin. 2009a. The extended prosodic word. In Kabak, Barıs & Grijzenhout, Janet (eds.), *Phonological Domains. Universals and Deviations*, 135–194. New York: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110217100.2.135

Ito, Junko & Mester, Armin. 2009b. The onset of the prosodic word. In Parker, S. (ed.), *Phonological Argumentation. Essays on Evidence and Motivation*, 227–260. London: Equinox. DOI: https://doi.org/10.1558/equinox.29400

Ito, Junko & Mester, Armin. 2019. Match as Syntax-Prosody MAX/DEP: Prosodic Enclisis in English. *English Linguisticsi* 36(1). 1–28. DOI: https://doi.org/10.9793/elsj.36.1_1

Johnson, Kyle. 1997. *Bridging the gap. In search of the English middle field.* Amherst, MA: University of Massachusetts manuscript.

Jokweni, Mbulelo W. 1995. *Aspects of IsiXhosa phrasal phonology.* Urbana Champaign: University of Illinois dissertation.

Kaisse, Ellen M. 1983. The syntax of auxiliary reduction in English. *Language* 59. 93–122. DOI: https://doi.org/10.2307/414062

Kallulli, Dalina. 2006. Triggering factivity: Prosodic evidence for syntactic structure. In Baumer, Donald & Montero, David & Scanlon, Michael (eds.), *Proceedings of 25th West Coast Conference on Formal Linguistics*, 211–219. Sommerville, MA: Cascadilla Proceedings Project.

Kaltenböck, Gunther. 2007. Spoken parenthetical clauses in English: A taxonomy. In Dehé, Nicole & Kavalova, Yordanka (eds.), *Parentheticals*, 25–52. Amsterdam/Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/la.106.05kal

Kan, Seda. 2009. *Prosodic domains and the syntax-prosody mapping in Turkish*. Boğaziçi University dissertation.

Kandybowicz, Jason. 2007. Comp-trace effects explained away. In Baumer, Donald & Montero, David & Scanlon, Michael (eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics*, 220–228. Somerville, MA: Cascadilla.

Kearns, Kate. 2007. Regional variation in the syntactic distribution of null finite complementizer. *Language Variation and Change* 19. 295–336. DOI: https://doi.org/10.1017/S0954394507000117

Keating, Patricia. 2003. Phonetic encoding of prosodic structure. In Palethorpe, Sallyanne & Tabain, Marija (eds.), *Proceedings of the 6th International Seminar on Speech Production,* 119–124. Sydney, Australia.

Keating, Patricia, & Cho, Taehong & Fougeron, Cécile & Hsu, Chai-Shune. 2004. Domain-initial strengthening in four languages. In Local, John & Ogden, Richard & Temple, Rosalind (eds.), *Papers in Laboratory Phonology VI*, 143–161. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511486425.009

King, Benjamin F. 1906. *Ben King's Verse* (N. Waterman, ed.). Chicago: Forbes & Company. (Original work published in 1898).

Kuno, Susumu. 1976. Gapping: A Functional Analysis. *Linguistic Inquiry* 7(2). 300–318.

Labov, William. 1969. Contraction, Deletion, and Inherent Variability of the English Copula. *Language,* 45(4). 715–762. DOI: https://doi.org/10.2307/412333

Ladd, Robert. 1986. Intonational phrasing: The case for recursive prosodic structure. *Phonology* 3. 311–340. DOI: https://doi.org/10.1017/S0952675700000671

Ladd, Robert. 1996. *Intonational Phonology*. Cambridge, UK: Cambridge University Press.

Lahiri, Aditi & Plank, Frans. 2010. Phonological phrasing in Germanic: the judgement of history, confirmed through experiment. *Transactions of the Philological Society* 108(3). 370–398. DOI: https://doi.org/10.1111/j.1467-968X.2010.01246.x

Lass, Roger & Anderson, John M. 1975. *Old English phonology*. Cambridge: Cambridge University Press.

Liang, Yiming & Amsili, Pascal & Burnett, Heather. 2021. New ways of analyzing complementizer drop in Montréal French: Exploration of cognitive factors. *Language Variation and Change* 33. 359–385. DOI: https://doi.org/10.1017/S0954394521000223

Lohndal, Terje & Samuels, Bridget. 2013. Linearizing Empty Edges. In Folli, Raffaella & Sevdali, Christina & Truswell, Robert (eds.), *Syntax and its Limits*, 66–79. Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199683239.003.0004

Marwick, James D. (ed.). 1869. *Extracts From the Records of the Burgh of Edinburgh, 1403–1528*. Edinburgh: Scottish Burgh Records Society.

McCawley, James D. 1982. Parentheticals and discontinuous constituent structure. *Linguistic Inquiry* 13(1). 91–106.

McElhinny, Bonnie S. 1993. Copula and Auxiliary Contraction in the Speech of White Americans. *American Speech* 68(4). 371–399. DOI: https://doi.org/10.2307/455773

Mindt, Ilka. 2011. *Adjective Complementation: An empirical analysis of adjectives followed by* that-*clauses*. Amsterdam/Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/scl.42

Mott, Edward Harold. 1895. *The Old Settler, the Squire and Little Peleg*. New York: United States Book Company.

Murray, James, A. H. 1873. *The Dialect of the Southern Counties of Scotland*. London: Asher & Co.

Myrberg, Sara. 2010. *The intonational phonology of Stockholm Swedish*. Stockholm University dissertation.

Myrberg, Sara. 2013. EQUALSISTERS and prosodic branching. *Phonology* 30(1). 73–124. DOI: https://doi.org/10.1017/S0952675713000043

Napoli, Donna Jo. 1982. Initial material deletion in English. *Glossa* 16(1). 85–111.

Nespor, Marina & Vogel, Irene. 1986. *Prosodic Phonology*. Dordrecht: Foris.

Onaka, Akiko. 2003. Domain-initial strengthening in Japanese: An acoustic and articulatory study. In Recasens, Daniel & Solé, Maria-Josep (eds.), *Proceedings of the 15th International Congress of Phonetic Sciences,* 2091–2094. Barcelona, Spain: ICPhS Organizing Committee.

Ormazabal, Javier. 1995. *The syntax of complementation: On the connection between syntactic structure and selection*. Storrs, CT: University of Connecticut dissertation.

Pak, Marjorie. 2008. *The postsyntactic derivation and its phonological reflexes*. Philadelphia, PA: University of Pennsylvania dissertation.

Paster, Mary Elizabeth. 2006. *Phonological Conditions on Affixation*. Berkeley, CA: University of California dissertation.

Perlmutter, David M. 1968. *Deep and surface structure constraints in syntax*. New York: Holt, Rinehart and Winston.

Pesetsky, David. 1995. *Zero syntax: Experiencers and Cascades*. Cambridge, MA: MIT Press. (Original manuscript circulated in 1992).

Pesetsky, David & Torrego, Esther. 2001. T-to-C Movement: Causes and Consequences. In Kenstowicz, Michael (ed.) *Ken Hale: A Life in Language*, 355–426. Cambridge: MIT Press. DOI: https://doi.org/10.7551/mitpress/4056.003.0014

Pierrehumbert, Janet B. 1980. *The phonology and phonetics of English intonation.* Cambridge, MA: MIT dissertation.

Pierrehumbert, Janet B. & Talkin, David. 1992. Lenition of /h/ and glottal stop. In Docherty, Gerry & Ladd, Robert (eds.), *Gesture, segment, prosody (Papers in Laboratory Phonology 2)*, 90–117. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511519918.005

Polsue, Joseph. 1867. *A Complete Parochial History of the County of Cornwall.* Truro: W. Lake.

Postal, Paul. 1974. *On raising: One rule of English grammar and its theoretical implications.* Cambridge, MA: MIT Press.

Potts, Christopher. 2004. *The logic of conventional implicatures.* Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199273829.001.0001

Quirk, Randolph & Greenbaum, Sidney & Leech, Geoffrey & Svarvik, Jan. 1972. *A grammar of contemporary English.* New York: Harcourt Brace Jovanovich, Inc.

Quirk, Randolph & Greenbaum, Sidney & Leech, Geoffrey & Svartvik, Jan. 1982. *A University Grammar of English*. London: Longman.

Quirk, Randolph & Greenbaum, Sidney & Leech, Geoffrey & Svartvik, Jan. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

Richards, Norvin. 1999. Complementizer cliticization in Tagalog and English. In Smallwood, Catherine & Kitto, Carolyn (eds.), *Toronto working papers in linguistics 16(2)* [*Proceedings of AFLA 6*], 297–312. Toronto: University of Toronto.

Rissanen, Matti. 1991. On the history of *that*/zero as object clause links in English. In Altenberg, Bengt & Aijmer, Karin (eds.), *English Corpus Linguistics: Studies in Honour of Jan Svartvik*, 272–289. Longman.

Rosenbaum, Peter S. 1965. *The grammar of English predicate complement constructions.* Cambridge, MA: MIT dissertation.

Ross, John Robert. 1967. *Constraints on variables in syntax.* Cambridge, MA: MIT dissertation.

Sag, Ivan A. 1980. *Deletion and Logical Form.* New York: Garland Publishing.

Schütze, Carson T. 1994. Serbo-Croatian second position clitic placement and the phonology-syntax interface. In Carnie, Andrew & Harley, Heidi & Bures, Tony (eds.), *MIT Working Papers in Linguistics 21, Papers on phonology and morphology*, 373–473. Cambridge, MA: MIT, Department of Linguistics and Philosophy, MITWPL.

Selkirk, Elisabeth. 1972. *The phrase phonology of English and French.* Cambridge, MA: MIT dissertation. (Published 1980, New York: Garland).

Selkirk, Elisabeth. 1978. On prosodic structure and its relation to syntactic structure. In Fretheim, Thorstein (ed.), *Nordic Prosody II: Papers from a symposium*, 111–140. Trondheim: TAPIR.

Selkirk, Elisabeth. 1984. *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT Press.

Selkirk, Elisabeth. 1996. The prosodic structure of function words. In Morgan, James L. & Demuth, Katherine (eds.) *Signal to Syntax: Bootstrapping From Speech To Grammar in Early Acquisition*, 187–214. Mahwah, NJ: Lawrence Erlbaum.

Selkirk, Elisabeth. 2002. Contrastive focus vs. presentational focus: Prosodic evidence from right node raising in English. In Bel, Bernard & Marlien, Isabelle (eds.), *Speech Prosody 2002: Proceedings of the 1st International Conference on Speech Prosody,* 643–646. Aix-en-Provence.

Selkirk, Elisabeth. 2005. Comments on intonational phrasing in English. In Frota, Sónia & M. Vigário, Marina & Freitas, M. João (eds.), *Prosodies*, 11–58. Berlin: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110197587.1.11

Selkirk, Elisabeth. 2009. On clause and intonational phrase in Japanese: The syntactic grounding of prosodic constituent structure. *Gengo Kenkyu (Journal of the Linguistic Society of Japan)* 136. 35–73.

Selkirk, Elisabeth. 2011. The syntax-phonology interface. In Goldsmith, John & Riggle, Jason & Yu, Alan C. L. (eds.), *The Handbook of Phonological Theory*, 435–483. Oxford: Wiley-Blackwell. DOI: https://doi.org/10.1002/9781444343069.ch14

Shank, Christopher C. & Van Bogaert, Julie & Plevoets, Koen. 2016. The diachronic development of zero complementation: A multifactorial analysis of the that/zero alternation with think, suppose, and believe. *Corpus Linguistics and Linguistic Theory* 12(1). 31–72. DOI: https://doi.org/10.1515/cllt-2015-0074

Stanton, Juliet. 2021. Allomorph selection precedes phonology: Evidence from Yindjibarndi. *Natural Language & Linguistic Theory* 40. 1317–1352. DOI: https://doi.org/10.1007/s11049-021-09531-0

Stowell, Tim. 1981. *Origins of phrase structure.* Cambridge, MA: MIT dissertation.

Suárez-Gómez, Cristina. 2000. *That*/zero variation in private letters and drama (1420–1710). A corpus-based approach. *Miscelanea. A Journal of English and American Studies* 21. 179–204. DOI: https://doi.org/10.1515/jhsl-2015-0003

Swingle, Kari. 1993. The role of prosody in right node raising. In Pullum, Geoffrey & Potsdam, Eric (eds.), *Syntax at Santa Cruz 2*, 82–112. UC Santa Cruz.

Tabain, Marija. 2003. Effects of prosodic boundary on /aC/ sequences: articulatory results. *Journal of the Acoustical Society of America* 113. 2834–2849. DOI: https://doi.org/10.1121/1.1564013

Tabain, Marija & Perrier, Pascal & Savariaux, Christophe. 2003. A kinematic study of prosodic boundary effects on /i/ articulation in French. In Recasens, Daniel & Solé, Maria-Josep (eds.), *Proceedings of the 15th International Congress of Phonetic Sciences,* 2617–2620. Barcelona, Spain: ICPhS Organizing Committee.

Truckenbrodt, Hubert. 1995. Extraposition from NP and prosodic structure. *North East Linguistics Society* 25(1). 503–517.

Truckenbrodt, Hubert. 1999. On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30(2). 219–255. DOI: https://doi.org/10.1162/002438999554048

Truckenbrodt, Hubert. 2005. A short report on intonation phrase boundaries in German. *Linguistische Berichte* 203. 273–296.

Truckenbrodt, Hubert. 2015. Intonation phrases and speech acts. In Kluck, Marlies & Ott, Dennis & de Vries, Mark (eds.), *Parenthesis and ellipsis. Crosslinguistic and theoretical perspectives*, 301–349. Berlin: Walter de Gruyter. DOI: https://doi.org/10.1515/9781614514831.301

Tyler, Matthew. 2019. Simplifying Match Word: Evidence from English functional categories. *Glossa: a journal of general linguistics* 4(1):15. 1–32, DOI: https://doi.org/10.5334/gjgl.631

Van Luven, Katie. 2018. *Pseudoclefts*. Ottawa, Ont.: Carleton University dissertation. DOI: https://doi.org/10.22215/etd/2018-13216

Webelhuth, Gert. 1992. *Principles and parameters of syntactic saturation.* New York: Oxford University Press.

Wheeldon, Linda & Lahiri, Aditi. 1997. Prosodic units in speech production. *Journal of Memory and Language* 37. 356–381. DOI: https://doi.org/10.1006/jmla.1997.2517

White, Laurence & Turk, Alice E. 2010. English words on the Procrustean bed: Polysyllabic shortening reconsidered. *Journal of Phonetics* 38. 459–471. DOI: https://doi.org/10.1016/j.wocn.2010.05.002

Wright, Joseph. 1903. *AT*. In Joseph Wright (ed.), *Volume 1 of The English Dialect Dictionary: Being the Complete Vocabulary of All Dialect Words Still in Use, Or Known to Have Been in Use During the Last Two Hundred Years; Founded on the Publications of the English Dialect Society and on a Large Amount of Material Never Before Printed,* 87. New York: H. Frowde.

Zec, Draga & Inkelas, Sharon. 1990. Prosodically constrained syntax. In Inkelas, Sharon & Zec, Draga (eds.), *The Phonology-Syntax Connection*, 365–378. Chicago: University of Chicago Press.

Zwicky, Arnold. M. 1970. Auxiliary Reduction in English. *Linguistic Inquiry* 3(1). 323–336.

Zwicky, Arnold. M. 1971. In a Manner of Speaking. *Linguistic Inquiry* 2(2). 223–233.

Zwicky, Arnold. M. & Pullum, Geoffrey. K. 1983. Cliticization vs. Inflection: English N'T. *Language* 59(3). 502–513. DOI: https://doi.org/10.2307/413900