19 Shaping Rhythm to Keep Balance: The Structural Implications of Temporal Modulation

NINO GRILLO, ANDREA SANTI, AND GIUSEPPINA TURCO

19.1 Introduction

One fundamental constraint on natural language is that, while linguistic representations are hierarchical, words have to be produced (and heard) one after the other. Many cognitive scientists argue this requires that speakers and hearers map the hierarchical information associated with sentences and their meanings into (and from) linearized strings of linguistic material while minimizing ambiguity and information loss. This mapping is made possible by the complex interaction of word order, morphology, and prosody. At the prosodic level, duration (the focus of this chapter), together with pitch and intensity, contributes both to speech segmentation (e.g., to identify word boundaries) and to grouping words into hierarchically organized prosodic representations (partially) mapping into underlying syntactic structures and their semantic interpretation.

Our ability to use durational information for speech disambiguation depends in turn on the ability to track the rhythmic properties of speech, as defined in this volume. For example, while speech rate itself does not carry grammatical information, it largely shapes our perception of both segmental and suprasegmental features, leading us to perceive the same speech sounds as being longer in the context of faster speech rate and shorter in the context of slower speech rate.[1] In other words, speech rate provides a baseline against which we can evaluate the durational properties of segmental and suprasegmental information against our predictions of their expected duration at that specific rate. At segmental level, effects of rate normalization have been reported for, as examples, voice onset time (Miller and Grosjean, 1981; and others), vowel length (Reinisch and Sjerps, 2013; and others), and word segmentation (Salverda et al., 2003; Reinisch et al., 2011). Similarly, in the domain of phrasal prosody, the realization and interpretation of prosodic boundaries have also been shown to be contextually dependent in production and their interpretation in parsing (on the importance of relative versus absolute

[1] See, for example, the discussion of *rate normalization* in Chapter 5 and references cited therein.

https://doi.org/10.1017/9781009295888.023 Published online by Cambridge University Press

values in the perception of boundaries, see, for example, Schafer, 1997; Carlson et al., 2001; Clifton Jr et al., 2002; Frazier et al., 2006; and others).

The durational properties of speech, however, are not only determined by structural factors but can also be modeled as a function of predictability: Less predictable words and segments are typically produced more slowly, and more predictable material is produced more quickly. Considerations of predictability and structure often make aligned predictions in relation to duration (e.g., in the domain of focus or in the case of early closure discussed below). To distinguish the relative contribution of these two factors, we identified a case in which the two factors make opposite predictions: the contrast between structural sisterhood (e.g., *The horse raced past the barn and fell*) and nesting (e.g., *The horse raced past the barn fell*). It is well established that nested structures are harder to process and understand (at least out of the blue) than sisterhood structures. While the source of the so-called garden path effects triggered by nesting is still a contentious matter, predictability-based accounts have become increasingly more prominent in psycholinguistics (Hale, 2001; Levy, 2008).

The main aim of this chapter is to argue that there are well-defined sets of environments for which lower predictability is in fact associated with shorter duration due to prosodic principles taking precedence over predictability in shaping the durational properties of speech. After briefly summarizing the role of prosodic structure and predictability in duration, we report on four recent studies showing that an inverse relation between duration and complexity is observed in the prosodic disambiguation of nested garden path sentences in production and comprehension. Specifically, the results of these studies show that, across different syntactic categories, nested garden path sentences are exceptional in displaying a relatively faster tempo when disambiguated towards less predictable structures. The important conclusion here is that, while predictability plays an important role in modulating duration, when in conflict with prosody, prosodic structure takes precedence.

We see the acceleration of a speech rhythm baseline as a hallmark of structural nesting in garden path sentences, and we suggest an account in which this change of tempo stems from an interaction between two types of principles: principles governing the syntax–prosody mapping, and principles balancing the size of prosodic phrases.[2] Nesting naturally generates a conflict between these two types of principles: Syntactically and semantically, it creates complex objects that should ideally be prosodified as a single constituent. This, however, would result in longer phrases, which can be difficult to manage for balancing principles. One solution is to break these long syntactic objects into separate prosodic phrases (as in extraposition; Wagner, 2005, 2010). We argue that nested garden path sentences, however, are unique in preventing nested elements from being prosodically separated from their hosts. In these structures the need for balance is satisfied by an increasing tempo of the long phrase. This solution increases parsability of nested garden paths while satisfying both types of principles. In essence, the prosodic pattern of nested garden path sentences exemplifies how different levels of representation interact and impact the rhythmic properties of speech. In simpler terms, disrupting the rhythm of speech at a lower level (tempo acceleration) helps create a smoother rhythm at a higher level (balanced prosodic phrases).

[2] Principles of balance might be argued to originate from neural oscillatory mechanisms, discussed in more detail in Chapter 18.

https://doi.org/10.1017/9781009295888.023 Published online by Cambridge University Press

19.1.1 Duration and Prosodic Structure

We will follow the common assumption that prosodic structures constitute an independent level of representation. In this perspective, prosody can be understood as the structure that determines and organizes the acoustic realization of an utterance in relation to its phrasing (the chunking of speech) and (lexical and post-lexical) prominence configuration, among other aspects of speech (i.e., voice quality, rate; Selkirk, 1980; Beckman and Pierrehumbert, 1986; Nespor and Vogel, 2007). Prosodic structure is conceived as a grammatical system, made of hierarchically ordered constituents, that is distinct from syntactic structure (but see other approaches to the prosody–syntax interface; Cooper and Paccia-Cooper, 1980; Gee and Grosjean, 1983) and is universal (with the implementation of language-specific adjustments). Figure 19.1 schematizes prominence and phrasing-related constituents of the prosodic structure as posited by well-known phonological accounts of prosody (i.e., the autosegmental-metrical [AM] framework; Pierrehumbert, 1980; Beckman and Pierrehumbert, 1986; Ladd, 1986) and described using notational conventions such as the ToBI (Tones and Break Indices) transcription system (Silverman et al., 1992; Beckman and Ayers, 1997).

While syntax and prosodic structure are not isomorphic, the relative degree of structural and interpretive integration of linguistic elements is one strong determinant of an utterance's prosodic structure. Principles governing the interface of prosody and syntax, such as edge alignment (Selkirk, 2000) and wrap (Truckenbrodt, 1995), push for prosodic phrasing to align with syntactic phrases as much as possible.

Box 19.1 Syntax–Prosody Mapping

Edge alignment:

"The right edge of any syntactic phrase (XP) in syntactic structure must be aligned with the right edge of a Major Prosodic Phrase in prosodic structure." (Selkirk, 2000, p. 232)

Wrap:

"Each syntactic XP must be contained in a phonological phrase." (Truckenbrodt, 1995, p. 10)

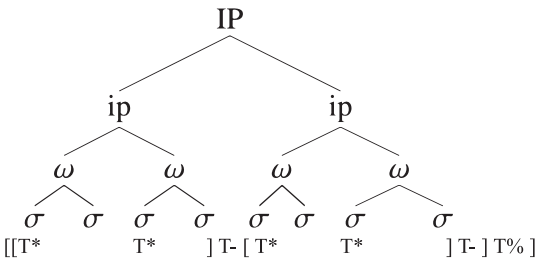


Figure 19.1 Prosodic hierarchy.
Schematic representation of the prosodic hierarchy. "IP" stands for intonational phrase (demarcated by boundary tones "T %") and "ip" for intermediate phrase (demarcated by phrase accents "T-") grouping words ("ω") and syllables ("σ"). T* stand for pitch accents realized on lexically stressed syllables.

https://doi.org/10.1017/9781009295888.023 Published online by Cambridge University Press

Within the realm of duration, pre-boundary lengthening (for English; Klatt, 1976; Wightman et al., 1992), pauses (Watson and Gibson, 2005; Breen et al., 2011), and domain-initial strengthening (Cho and Keating, 2009) are the main phenomena in which duration determines whether a boundary is perceived and what strength is associated with it (see the discussion in Wagner and Watson, 2010; Turk and Shattuck-Hufnagel, 2014; Dahan, 2015, and references cited therein).

For our purposes, a good illustration of how duration guides prosodic grouping at the phrasal level across both production and comprehension comes from Kjelgaard and Speer's (1999) work on the disambiguation of garden path sentences such as (1). In (1a) (i.e., the most accessible interpretation in the absence of prosodic disambiguation), the verb *leaves* is interpreted transitively and forms a single phrase with the following determiner phrase (DP) *the house*, which is followed by a prosodic break (indicated with "‖" here). In (1b) the same verb is interpreted intransitively, and it constitutes the final region of the temporal modifier *when John leaves*. This region is immediately followed by a prosodic boundary that signals that the following DP *the house* should not be integrated in the same phrase but should be interpreted as the subject of the following predicate *is dark*. This parse is strongly dispreferred in the absence of prosodic information or punctuation. The preference for interpreting the verb transitively (as in 1a) can alternatively be interpreted as part of a generalized parsing preference to incorporate incoming words into the phrase being processed whenever grammatically possible (*late closure strategy*; Frazier, 1979) or as generated by the relative higher frequency of transitive versus intransitive readings of verbs such as *leaves* (Tanenhaus et al., 1989; MacDonald et al., 1994).

(1) a. When John leaves the house ‖ it's dark. Late closure
b. When John leaves ‖ the house is dark. Early closure

Kjelgaard and Speer show that lengthening of the pre-boundary word *leaves* in (1b), compared to when the same word occurs in phrase-medial position in the late closure example in (1a), helps listeners avoid a garden path effect, even in the absence of an actual pause between the verb and the following DP. It is important to stress that what is critical for disambiguation is information about *relative* duration: The presence of a boundary in a given region will be a matter of relative duration of a given segment, not of the absolute durational properties of that segment. For a detailed discussion of how the importance of relative duration and global, rather than local, measures support a role for prosodic representations, see, for example, Schafer (1997); Carlson et al. (2001); Clifton Jr et al. (2002); Frazier et al. (2006); Shatzman and McQueen (2006); Speer and Blodgett (2006).

Constraints favoring the alignment of syntactic and prosodic phrasing, however, operate in tandem with other eurythmic[3] or balance constraints, which ensure, for example, that prosodic phrases do not exceed a certain size in production, and are preferably parsed into units of similar length in comprehension (Frazier and Fodor, 1978; Gee and Grosjean, 1983; Ghini, 1993; Fodor, 2002).

[3] Eurhythmic: well balanced, proportional, harmonious. In the domain of sentence processing, Frazier and Fodor's (1978) *sausage machine* model was an early implementation of the idea that sentences are parsed into prosodic phrases of approximately the same length.

https://doi.org/10.1017/9781009295888.023 Published online by Cambridge University Press

Box 19.2 Constraints on Length of Prosodic Phrases

Binary minimum:

"A major phrase must consist of at least two minor/accentual phrases." (Selkirk, 2000)

Binary maximum:

"A major phrase may consist of at most two minor/accentual phrases." (Selkirk, 2000)

Uniformity:

"A string is ideally parsed into same length units." (Ghini, 1993)

Conflicts can arise between these different constraints and generate mismatches between syntactic and prosodic phrasing such as the one in (2) and (3) (adapted from Turk and Shattuck-Hufnagel, 2014, p. 2):

(2) Syntactic structure:
[This is [the cat [that ate [the rat [that ate the cheese]]]]]

(3) Prosodic structure:
[This is the cat] [that ate the rat] [that ate the cheese].

The extent of syntax–prosody mismatch is rather controversial (for a recent summary, see Bennett and Elfner, 2019) and outside the immediate scope of this chapter, so we will not devote a long discussion to it. But to appreciate the debate, note that some researchers (e.g., Wagner, 2005, 2010) suggest that true instances of a mismatch might be more limited than standardly assumed, and that cases such as (2) and (3) can be accounted for by extraposition of the relative clauses (RCs), such that they attach to a higher position in the syntactic structure and provide a much tighter alignment between prosodic and syntactic structures. Whatever the correct account for these and other similar contrasts, they serve well to illustrate the tension between syntax–prosody mapping and eurhythmic principles, an issue we'll return to below when discussing the prosodic properties of nested garden path sentences.

19.1.2 Duration and Predictability

Durational properties of speech have also been shown to be largely dependent on predictability. More predictable words/segments are reliably associated with a shorter production duration. Less predictable elements of utterances, on the other hand, are more carefully articulated, and thus produced more slowly (Lieberman, 1963; Aylett, 2000; Jurafsky et al., 2001; Aylett and Turk, 2004; Turk and Shattuck-Hufnagel, 2014; Levy and Jaeger, 2007; and others).

For example, syllables in words that carry new or contrasting information, and are thus less predictable, are typically associated with prosodic prominence, making them louder, longer,

https://doi.org/10.1017/9781009295888.023 Published online by Cambridge University Press

and articulated more carefully (Aylett, 2000; Aylett and Turk, 2004; Watson et al., 2008; Turk and Shattuck-Hufnagel, 2014).[4] Aylett (2000) and Aylett and Turk (2004) suggest an information theoretic account of these observations, the *smooth signal redundancy hypothesis*, according to which the inverse relation between duration and predictability "provides an efficient way of ensuring that elements with low levels of language redundancy are produced for a longer period of time and perhaps with more salient acoustic characteristics, and will thus be likely to be recognized" Aylett and Turk (2004, p. 33). This is illustrated in (4), where the words *pizza* and *John* are associated with longer duration when in focus. It is also well known that elements in prosodically prominent positions receive more attention, are remembered better, and are processed faster than non-prominent elements (Cutler, 1976, 2012).[5]

(4) a. What does John like? John likes PIZZA.
b. Who likes pizza? JOHN likes pizza.

Similar considerations apply to the domain of syntactic phrasing, where more predictable parses are expected to be produced faster than less predictable ones. The early closure garden path sentence discussed above (and repeated in 5) provides a good illustration of this. As mentioned, according to one analysis of early closure, the intransitive reading in (5b) is harder to parse because it is less predictable, for example, because it is encountered less frequently, than the transitive reading in (5a) (Tanenhaus et al., 1989; MacDonald et al., 1994). A predictability account would assume that a longer duration of the ambiguous region (*leaves the house*) would help listeners to resolve the ambiguity towards the less predictable parse (presumably by positing a boundary between *leaves* and *the house*).

(5) a. When John leaves the house || it's dark.
b. When John leaves || the house is dark.

As this example illustrates, the same durational difference could in principle be correctly predicted on the basis of both prosodic structure and predictability (i.e., longer duration at *leaves* for early closure than late closure). In many instances of ambiguity, it would indeed seem that the effects on production of both prosodic structure and predictability are essentially aligned: "when predictability is low, syllables are more likely to be prosodically prominent, and words are more likely to be demarcated using prosodic boundary correlates such as initial- and final-lengthening and pause" (Turk and Shattuck-Hufnagel, 2014, p. 4). However, as Turk and Shattuck-Hufnagel (2014, p. 4) also point out, "the problem with predictability as a factor affecting duration is that it is unclear whether prosodic structure and predictability are both motivated as separate and independent factors affecting duration. This is because prosodic structure and predictability are not independent."

To dissociate the relative contribution of predictability and prosodic structure, it is important, therefore, to focus on cases in which these two factors make opposite predictions. In recent

[4] On the effect of predictability on syllable duration and pronunciation, see also Greenberg (1999, 2006); Greenberg et al. (2003).

[5] Notice, however, that, as Cutler and Fodor (1979) showed, the main determinant here is not acoustic prominence per se but the grammatical category of focus.

https://doi.org/10.1017/9781009295888.023 Published online by Cambridge University Press

work (Grillo and Turco, 2016; Grillo et al., 2018, 2019), we have suggested that the contrast between sisterhood and nesting is one such case.

19.2 Prosody and Predictability Make Opposite Predictions for Nested Garden Paths

The structural ambiguity of nesting versus sisterhood is extremely common and has been one of the main testing grounds for different theories of sentence processing. The sentences in (6) are a few examples of well-described cases of locally ambiguous sentences involving this structural ambiguity. As shown by the bracketing, the first example of each pair of sentences illustrates *nesting*: The ambiguous phrase is embedded within the constituent it modifies and forms a single syntactic phrase with it. The second example of each sentence pair illustrates *sisterhood*: In these examples the ambiguous phrase is a sister constituent to the preceding constituent. For each set of examples, the nested reading has been shown to generate longer reading times at the disambiguation region and poorer comprehension (at least out of the blue) than the sisterhood reading (see, for example, Pickering and Van Gompel, 2006, for a review).[6]

(6)	a.	[[The horse [raced past the barn]] fell]	Reduced relative
	b.	[[The horse] [raced past the barn] and [fell]]	Main verb
	c.	John [told [the man [that he was running with]] [to wait]]	RC
	d.	John [told [the man] [that he was running with Max]]	CC
	e.	[Put [the horse in the barn] [on the truck]]	Restrictive PP
	f.	[Put [the horse] [in the barn]]	Goal PP
	g.	John [saw [the man [with the binoculars]]]	Restrictive PP
	h.	John [saw [the man] [with the binoculars]]	Instrumental PP

While the source of this contrast in processing difficulty is still a contentious matter, an increasingly prominent perspective in psycholinguistics is that nested structures are harder to parse because they are less predictable. This approach is supported by corpus studies and computational modeling. Hale (2001), for example, shows that the reduced relative parse in (6a) is seven times less likely to occur in a corpus than the simpler analysis involving the unmodified noun phrase. Similarly, Jurafsky (1996), based on data from Connine et al. (1984), combined syntactic probability of the main verb and reduced-RC parse (and the related lexical probability of the intransitive versus transitive reading of the verb *raced* in the *horse raced past*

[6] That the restrictive analyses in (6) are harder to parse out of the blue is undisputed; the theoretical debate has focused on the underlying causes of the garden path effect they generate and their relevance for sentence processing models. A variety of factors (including frequency of lexical frames and syntactic construction, besides semantic, pragmatic, and contextual factors) has in fact been argued to modulate the strength of the garden path effect (see, among many others, Bever, 1970; Kimball, 1973; Crain and Steedman, 1985; Frazier, 1987; Altmann and Steedman, 1988; Ferreira and Clifton, 1986; Trueswell et al., 1994; Jurafsky, 1996; Ni et al., 1996; McRae et al., 1998; Paterson et al., 1999; Hale, 2001; Sedivy, 2002; Clifton et al., 2003; Staub et al., 2018).

https://doi.org/10.1017/9781009295888.023 Published online by Cambridge University Press

the barn (and) fell) to estimate "the probability ratio of the two analyses of pre-disambiguation context . . . as roughly 82:1" (Levy, 2011).

From an information theoretic perspective, therefore, longer reading times observed in processing studies for nested structures are seen as a function of lower predictability. As with production, studies on sentence processing show an inverse relationship between redundancy and duration. Structural analyses entertained at earlier regions of a sentence generate structural expectations about parts of the sentence that are yet to come (Konieczny, 2000; Hale, 2001; Lau et al., 2006; Staub and Clifton, 2006; Levy, 2008; Levy et al., 2012; Traxler, 2014; Kuperberg and Jaeger, 2015). The amount of time spent reading a particular region is a function of the strength of prior expectations.

If the durational properties of speech are primarily determined by predictability, we should expect the lower predictability of nested garden paths to lead to longer durations in production than string-identical sentences involving more predictable sisterhood structures. This perspective should generate similar predictions for comprehension as well: A longer produced duration of the ambiguous region should make it easier for listeners to parse structural nesting; a shorter produced duration, on the other hand, should aggravate the garden path effect as it would be taken to map onto the more predictable sisterhood structure.

Importantly, considerations of prosodic structure generate the opposite predictions in relation to both production and comprehension of nested garden paths, leading us to expect a shorter duration for nesting than sisterhood in production, which translate into facilitated processing, that is, a reduction in the garden path effect for nesting when comprehending an ambiguous phrase produced with a relatively shorter duration. These predictions are based on the relative degree of structural (and interpretive) integration of nesting and sisterhood and the constraints on prosodic structure introduced above.

Sister constituents are independent phrases and attach higher in the syntactic structure than nested constituents. This syntactic difference makes sister constituents more likely to be produced as separate prosodic phrases than nested constituents, which are contained within the XP they modify, and are thus more likely to be mapped onto a single prosodic phrase (following the principles of edge alignment and wrap). Indeed, it is well established that a higher attachment site correlates with separate phrasing, and given that prosodic phrasing modulates the durational properties of utterances in predictable ways (e.g., pre- and post-boundary lengthening), separate phrasing typically leads to longer durations for attachment to higher positions and shorter durations for more deeply embedded strings (Hirschberg and Avesani, 1997; Clifton Jr et al., 2002; Wagner and Watson, 2010; Poschmann and Wagner, 2015; Grillo and Turco, 2016; among many others).

Given the differences in attachment site of sister and nested constituents, we can predict that sister constituents are more likely to be produced as separate phrases than nested constituents. In turn, this leads us to predict a longer duration for sisterhood structures at the regions preceding and following the phrase boundary, despite their higher predictability.[7]

[7] To increase readability, we postpone a discussion of how mapping constraints interact with balance constraints in the domain of nested garden paths until Section 19.3. As will be made clear, this does not affect the current discussion, as similar predictions obtain when these constraints are taken into account.

https://doi.org/10.1017/9781009295888.023 Published online by Cambridge University Press

The example from nested garden paths involving prepositional phrases (PPs) in (6e) versus (6f), repeated below, illustrates these effects very appropriately. These sentences have been consistently shown to be prosodically disambiguated (Snedeker and Trueswell, 2003; Kraljic and Brennan, 2005; Schafer et al., 2005; Speer et al., 2011). Importantly, a stronger prosodic boundary between the DP (*the horse*) and the PP (*in the barn*) leads to a longer duration of both regions in the more predictable and easier to parse sisterhood condition (7a) than in the less predictable nested structure in (7b).

(7) a. [Put [the horse] [in the barn]].
b. [Put [the horse [in the barn]] [on the truck]].

While this prosodic pattern is well attested, no work to date to our knowledge has attempted to integrate these results within a predictability account of duration. More generally, despite the central role played by nested garden paths in guiding psycholinguistic theories, until recently there has been surprisingly little experimental evidence on how the contrasts in (6) are prosodically disambiguated in production and comprehension.[8] A comprehensive study of the prosody of sisterhood versus nesting garden path sentences is still lacking. This is probably also due to the widespread assumption that some of these ambiguities (e.g., main verb versus reduced relatives) are not prosodically disambiguated (see, for example, Fodor, 2002; Wagner and Watson, 2010).

19.2.1 Reduced-RC Garden Paths

Recent results from our research group support the prediction that nested garden path sentences are associated with a shorter duration than the sisterhood structure, despite nesting being more difficult to parse out of the blue (and, at least in some cases, less predictable because of their lower frequency). In this section we review these results, and in the following section we will review our theoretical interpretation of these durational effects. In Grillo et al. (2018) we investigated the prosodic disambiguation of classic garden path sentences that are temporarily ambiguous between a main verb and a reduced-RC verb (6a, b)]. This type of garden path is generated by the ambiguity between a past tense and a passive past participle reading of the verb *raced*. As mentioned, we predicted that a difference in attachment height of the verb phrase (VP) (*raced past the barn*) with respect to the subject DP (*the horse*) should lead to shorter duration for the less predictable reduced-RCs. As explained above, the rationale for these predictions comes from the simple observation that the ambiguous string forms a single constituent in the reduced-RC analysis, but two separate syntactic phrases in the main verb analysis.

We also suggested that prosodic disambiguation of this type of contrast might have previously gone unnoticed because these sentences were typically presented in isolation and started with the critical pre-boundary ambiguous region. This is a problem because, as mentioned above, the absolute durational properties of the ambiguous region are not informative in isolation, but only when evaluated in relation to a baseline tempo provided by linguistic

[8] Harris et al. (2016) previously investigated the prosody of reduced relatives but did not compare it with that of minimally different sentences involving the main verb parse.

https://doi.org/10.1017/9781009295888.023 Published online by Cambridge University Press

material preceding the ambiguous region itself (Schafer, 1997; Carlson et al., 2001; Clifton Jr et al., 2002; Frazier et al., 2006).

In Grillo et al. (2018) we used the Planned Production paradigm to test these hypotheses. In Planned Production, participants are instructed to silently scan the entire sentence before producing them naturally and fluently at normal speed. All stimulus sentences were preceded by a short introductory main clause (*Greg said that* in (8)) that was neutral with respect to the relevant disambiguation, but provided a baseline tempo against which the durational properties of the ambiguous region could be evaluated:[9]

(8) a. Main verb condition:
Greg said that [$_{TP}$ [$_{DP}$ the businessmen] [$_{VP}$ loaned money at low interest] and [$_{VP}$ were told to record their expenses]].
b. Reduced-RC condition:
Greg said that [$_{TP}$ [$_{DP}$ the [$_{NP}$ businessmen [$_{CP}$ loaned money at low interest] [$_{VP}$ were told to record their expenses]]].

The results showed that the contrast in (8) is prosodically disambiguated, and that this disambiguation can be detected as early as the subject DP (*the businessmen*). We found the ambiguous region (*the businessmen loaned money at low interest*) to be shorter in the less predictable RC condition (8a) than the more predictable main clause condition (8b). This difference was visible both in terms of absolute and relative duration, calculated as the ratio of the ambiguous region duration to that of the intro clause (*Greg said that*). Speech rate increased significantly more for the less predictable nested condition than for the sisterhood condition. These durational differences cannot be accounted for from a predictability perspective, as they contradict the predictions it would make. Instead, they reflect the distinct structural relation between the DP and the VP in these sentences.

In a follow-up forced-choice comprehension study (Grillo et al., 2022), one of the participants' recordings from the Planned Production study of Grillo et al. (2018) were used to test the extent to which listeners use these prosodic cues to overcome the garden path effect. The disambiguating region was removed from the recordings of both the main verb and the reduced relative conditions. To test the hypothesis that the presence of a baseline tempo is essential to the interpretation of the temporal properties of the ambiguous region (i.e., to the interpretation and disambiguation of the boundary between *the businessmen* and *loaned money at low interest* in (8)), we also manipulated whether an introductory clause was present (9a–b) or not (9c–d):

(9) a. Reduced-RC prosody – baseline:
Greg said that *the businessmen loaned money at low interest* …
b. Main verb prosody – baseline:
Greg said that *the businessmen loaned money at low interest* …
c. Reduced-RC prosody – no baseline:
The businessmen loaned money at low interest …
d. Main verb prosody – no baseline:
The businessmen loaned money at low interest …

[9] Notice that, for convenience, we follow the original paper in referring to "main verb" parse, even though in the materials used the relevant verb is part of the embedded clause.

https://doi.org/10.1017/9781009295888.023 Published online by Cambridge University Press

The same durational (and more generally prosodic) differences (in absolute terms) are present in both pairs of sentences; that is, identical recordings are used for condition A and C, on the one side, and for conditions B and D, on the other. The only difference between the two is the presence of an introductory sentence, providing a slower baseline tempo, which was removed from the recording in conditions C and D.

Participants listened to the cropped recording and then selected between two alternative continuations (main verb versus reduced-RC) that were written on the screen. In line with previous results from the garden path literature, a strong preference for the main verb continuation was found across all conditions. Main verb-compatible continuations constituted approximately 80% of choices in both full (baseline present) and cut (baseline absent) main verb conditions. The reduced relative continuation, nevertheless, was chosen twice as often when the RC prosody was preceded by an introductory clause to provide a baseline tempo. This is predicted by our approach since without a baseline tempo it should be impossible to decide whether the durational properties of the ambiguous region should be taken to indicate lengthening/lack of lengthening. A much more likely outcome is that they will be taken to reflect grammatically uninformative absolute speech rate. We thus correctly predict a smaller garden path effect when reduced-RCs are embedded within a main clause providing a baseline rate.[10]

19.2.2 CC/RC Ambiguity

More recently, Grillo et al. (2023) showed that speakers similarly use temporal cues to prosodically disambiguate the complement clause (CC) versus RC garden path (10). The complementizer phrase (CP) *that he was singing with* is initially temporarily ambiguous between being a modifier (i.e., an RC) (10a) or a co-argument (i.e., a CC) of the DP, *the editor* (10a).

(10) a. RC
The kind lyricist [$_{V'}$ told [$_{DP}$ the [$_{CP}$ [$_{NP}$ editor [$_{C'}$ that he was singing with]]]] [$_{CP}$ to listen]].
paraphrasable as: *what the kind lyricist told the editor that he was singing with was to listen*
b. CC
The kind lyricist [$_{VP}$ [$_{V'}$ told [$_{DP}$ the [$_{NP}$ editor]] [$_{CP}$ that he was singing with Lola]]].
paraphrasable as: *what the kind lyricist told the editor was that he was singing with Lola*

The results of the Planned Production study showed that the ambiguous region (*the editor that he was singing (with)*) is associated with faster tempo when nested (the RC in (10a)) than when in a sisterhood relation (the CC in (10b)). In line with what we observed with reduced-RCs, this pattern was not only visible in raw durations (in ms) but also, and more importantly, in the change in speech rate between the intro phrase (*the kind lyricist told*) and the ambiguous region (*the editor that he was singing with*). Predictably, this pattern reversed at the final word of the ambiguous region (the preposition *with*), which precedes a major prosodic boundary in the RC condition.

[10] In an ongoing follow-up study (Guo et al., in preparation), we are manipulating the temporal properties of baseline and ambiguous regions while keeping all other prosodic dimensions constant. Our prediction is that the relative tempo of the baseline and the intro should strongly influence listeners' disambiguation of nested garden paths.

https://doi.org/10.1017/9781009295888.023 Published online by Cambridge University Press

Notice that we do not mean to suggest that tempo modulation is solely responsible for disambiguation of these structures. As can be clearly seen in Figure 19.2 (C, D), clear differences in pitch and intensity also differentiate the two readings (which is also supported by a ToBI analysis presented in Grillo et al., 2023). What is important for the present discussion is that the prosodic structure of RCs leads to (among other things) shorter duration, despite their lower predictability.

In a follow-up comprehension study (Grillo et al., 2023), we showed that listeners are sensitive to these prosodic differences in guiding their interpretations. In a forced-choice task, participants heard sentence fragments that again excluded the disambiguating region (e.g., *The kind lyricist told the editor that he was singing with*) and then selected between printed CC (*Lola*) and RC (*to listen*) continuations. The target sentences contained a prosodic disambiguation that was consistent with the Planned Production study. The results showed a strong effect of prosody, with the selection of an RC continuation more than doubling following the RC prosody (57.5%) compared to following the CC prosody (25.5%). A post hoc analysis that removed eight outlier participants who *only* selected a CC continuation showed that the two prosodic forms were equally informative to the choice of continuation; that is, participants chose an RC completion after hearing the RC prosody (70% of target RC responses) in equal proportion to a CC continuation after hearing a CC prosody (70% of target CC responses).

Summarizing thus far, we have presented converging evidence for a durational contrast between sisterhood and nesting. In each instance, temporal modulation (faster tempo in comparison to a baseline earlier in the sentence) has been observed for the more complex and less predictable nested reading. While the realization of these structures also differs along other prosodic dimensions (i.e., tonal), this durational pattern is clear and is observed both in novel studies from our research group (on the main verb/reduced relative ambiguity, the CC/RC ambiguity, the pseudo-relative/RC ambiguity) and for previous results from the literature (restrictive/goal PPs). While more work is ongoing to determine the relative contribution of tempo as opposed to other prosodic dimensions, it is nevertheless striking that listeners are able to converge on the less predictable reading in the presence of increased tempo that entails a shorter time to process the ambiguity. This brings about a significant conclusion: Prosodic structure and predictability are not always aligned. Prosodic structure seems to determine the durational properties of speech above and beyond predictability. This does not mean that predictability does not itself independently modulate duration but that when the two are in conflict, prosodic structure takes precedence. In future work, we plan to investigate the modulatory role of predictability on duration by adding different measures of predictability to our model (e.g., relative frequency of past tense versus past participle of different predicates).

19.3 The Interaction of Mapping and Eurhythmic Constraints in Garden Path Prosody

Following syntax–prosody mapping principles, we have proposed that the relative degree of structural (and interpretive) integration of nesting and sisterhood leads to specific predictions

https://doi.org/10.1017/9781009295888.023 Published online by Cambridge University Press

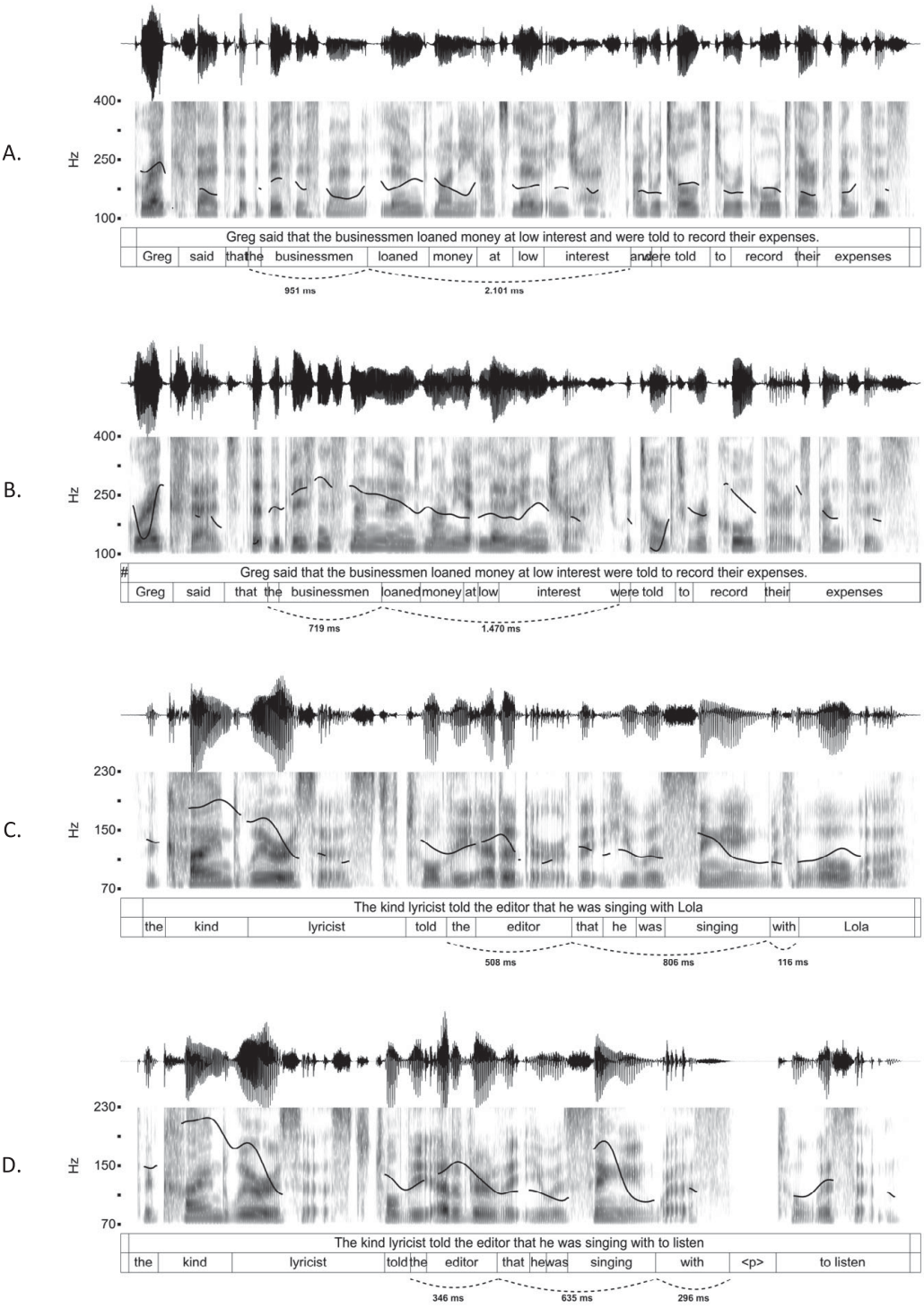


Figure 19.2 Spectrograms of nested garden path sentences.
Example of waveform, spectrogram, and fundamental frequency (F0) track of main verb (A), reduced-RC (B), CC (C), and RC (D) structures; recorded by British English native informants in Grillo et al. (2018, 2023) and used for the forced-choice experiments in Grillo et al. (2022, 2023) and segmented word by word. Absolute durations (in ms) of the nested material and the DP it modifies are indicated.

https://doi.org/10.1017/9781009295888.023 Published online by Cambridge University Press

about the durational properties of these structures. Sister constituents are by definition independent XPs and are thus more likely to be produced as separate prosodic phrases than nested constituents. Nested constituents, on the other hand, are contained within the XP they modify and are thus more likely to be mapped onto a single prosodic phrase (following the principles of edge alignment and wrap). This would correctly predict the pre- and post-boundary regions dictated by the relative strength of the boundary between sister and nested constituents.

As seen above, constraints favoring the alignment of syntactic and prosodic phrasing, however, are not absolute but operate in tandem with other eurhythmic or balancing constraints, which ensure that strings are ideally parsed into units of the same length and that do not exceed a certain size (Frazier and Fodor, 1978; Gee and Grosjean, 1983; Ghini, 1993; Selkirk, 2000; Fodor, 2002). Balance constraints lead a rhythmic pattern at the level of prosodic phrasing and might be rooted in more basic neural mechanisms (see Chapter 18). Nesting, and in particular nested garden paths, constitute a domain of natural tension between mapping and balancing constraints, providing a valuable foundation to evaluate their interaction. On the one hand, mapping constraints increase the likelihood that nested material is spelled out as a single constituent with the phrase it modifies. On the other hand, nesting increases the size of the host constituent, increasing the chances that the two will be split into separate phrases by balancing constraints. As we have seen above in Examples (1) and (2), this tension can result in separate phrasing in simpler cases, that is, when no garden paths are involved. Whether this happens through extraposition (which allows to also satisfy mapping constraints) or otherwise is out of the scope of this chapter. The crucial question we wish to address is, why does this not seem to happen in the domain of nested garden paths? Our answer is that separate phrasing would only worsen the garden path effect. We illustrate this claim through the familiar contrast between main verb and CCs (sisterhood) and (reduced-)RCs.

Prosodification of main verb and CCs in (11a, b) is straightforward: Principles of syntax–prosody mapping will lead to a preference to generate independent, and fairly balanced, phrases for the DP [*the horse*] and the VP [*raced past the barn*] and for the DP [*the woman*] and the CP [*that was running with Max*].

(11) a. The horse || raced past the barn || and fell.
b. John told the woman || that he was running with Max.

In the case of nesting, however, it is easy to see that a conflict will arise between these two types of principles. On the one hand, we can expect syntax–prosody mapping principles to push for the nested phrase to be produced as a single prosodic phrase with the head it modifies, as in (12a, b). This phrasing, however, will lead to a nonuniform pattern for the two sister constituents ([*the horse raced past the barn*] and [*fell*]; [*the woman that he was running with*] and [*to leave*]). Eurhythmic principles will resist this unbalanced (and heavy) prosodification, pushing for alternative phrasing, as, for example, the one in (12c, d).

(12) a. The horse raced past the barn || fell.
b. John told || the woman that he was running with || to wait.
c. The horse || raced past the barn || fell.
d. John told the woman || that he was running || with to wait.

https://doi.org/10.1017/9781009295888.023 Published online by Cambridge University Press

The problem with the prosodifications in (12c, d) is that the boundary between *the horse* and *raced* (and between *the woman* and *that*) actually encourages the incorrect parse, potentially generating an even stronger garden path effect. We follow the rational speaker hypothesis (Clifton Jr et al., 2002) that claims that speakers use prosody in an "internally consistent, rational, fashion, and that the listener assumes such rationality in interpretation" (Frazier et al., 2006, p. 246). Paraphrasing Frazier et al. in this perspective, if a speaker intends a structure where a constituent contains the reduced-RC, she will not insert a prosodic boundary that separates the reduced relative from the rest of its constituent without good reason. We thus expect that (reduced-)RCs in ambiguous environments will be produced as a single prosodic phrase with the DP they modify, that is, as in (12a, b). The same reasoning applies to the other instances of nested garden path ambiguities, and we will not repeat it for space reasons. We thus argue that in the case of nested garden paths, the conflict between syntactic and length constraints will not be resolved in favor of length balance.

Nevertheless, another option might be available for achieving some balance without strongly violating syntactic mapping constraints and going against the rational speaker hypothesis. This alternative, already envisaged in Ghini (1993) as an integral part of his Uniformity Principle, involves reducing the size of the offending phrase by compressing it, that is, increasing its tempo.[11, 12]

Box 19.3 Uniformity Definition

"A string is ideally parsed into units of the same length phrases. The average weight of the phrase depends on tempo: at an average rate of speech (moderato), a phrase contains two phonological words; the number of Ws within a phrase increases or decreases by one by speeding up or slowing down the rate of speech." (Ghini, 1993, p. 56)

This global effect (global because it applies to the whole ambiguous region and is not just localized at the boundaries) is exactly what we observed in Grillo et al. (2018, 2023). Once again, increased tempo can be grammatically relevant, but only when evaluated relative to a baseline tempo, which explains why reduced relative garden paths appear not to be prosodically disambiguated when presented in isolation. Without a baseline tempo, the increased tempo associated with nesting cannot be interpreted as acceleration, but only as fast speech rate, which (while important in a paralinguistic dimension) is not grammatically relevant.

[11] In Ghini's proposal, increasing tempo leads to a reduction of the number of prosodic words within a phrase because it forces lighter words (e.g., functional material) to cliticize on heavier words. Effectively, this allows us to achieve the same results (a lighter phrase) with different means (increased speech rate).

[12] As mentioned above, this type of constraint might be argued to stem from neural oscillatory mechanisms discussed in more detail in Chapter 18.

https://doi.org/10.1017/9781009295888.023 Published online by Cambridge University Press

That some form of compression is essential for successful prosodification (and parsing) of nested structures is also supported by recent results on the processing of multiple center embedding (multiple nesting to use the present terminology) discussed in Fodor et al. (2018). The groundbreaking work of Fodor et al. shows that the intelligibility *in reading* (i.e., in the absence of explicit prosody) of famously unparsable sentences with multiple center embeddings (*the mouse that the cat that the dog chased bit died*) increases considerably (in predictable ways) when the relative weight of each phrase is carefully manipulated to encourage phrasings that optimize the conflicting syntactic and length constraints:

(13) The French woman ‖ the man I met loves ‖ died.

Fodor et al. (2018) show that by lengthening the first DP and shortening the following material, they can encourage the multiple center embedding to be parsed in three optimal phrases, as in (13). The same manipulation of size leads not only to better prosodification in production but also to improved comprehension in silent reading. If our analysis is on the right track, we predict that similar results could also be achieved in less well-balanced sentences (in terms of length), by modulating the relative weight of each phrase through speech rate (i.e., slowing down tempo in the first phrase and accelerating in the second one).

In future work we also aim to investigate the feasibility of extending this account to the better-studied case of polysyllabic shortening, that is, the inverse relationship between the size of a constituent and the duration of its primary stressed syllable (see Lehiste, 1972; Nooteboom, 1997; Davis et al., 2002; Salverda et al., 2003; Shatzman and McQueen, 2006; White and Turk, 2010; and others).

More work is needed to evaluate this account against potential alternative analyses. One possible alternative explanation for these effects, discussed in Santi et al. (2019), is that nested material is produced faster to optimize processing of main clauses by reducing the temporal distance between, for example, the head of a subject (*the horse*) and the predicate of the main clause (*fell*). Grillo and Turco (2016), however, show that the same durational contrast between sisterhood and embedding is also observed when the nested material is right-branched and thus does not interfere with the processing of the main clause. Another alternative is that nested material might be produced faster because it involves old, backgrounded, or not-at-issue information. Although this explanation may seem attractive, it is not without its challenges. The first one is that, contrary to this common assumption, modifiers can in fact contain new information (e.g., *in my class there's [a student [who met the president]]*, where the whole complex nominal, including the modifier [*who met the president*], arguably carries new information). Another, potentially stronger, argument against this kind of analysis comes from a comparison of restrictive and appositive RCs. Appositive relatives (*John, who is a great guy, arrived yesterday*) are the textbook case of backgrounded/not-at-issue phrases. Contrary to restrictive relatives, they attach higher in the structure and crucially are associated with stronger prosodic boundaries, commas, intonation, and are produced more slowly than nested material (Poschmann and Wagner, 2015). Nevertheless, much more work is needed to investigate how information structure and constituent structure interact in shaping prosody in nested garden paths, but see Guo et al. (2023, 2024a, 2024b, 2025) for some preliminary results in this domain.

https://doi.org/10.1017/9781009295888.023 Published online by Cambridge University Press

19.4 Conclusions

A number of recent studies from comprehension and production show that nested versus sisterhood structures are prosodically disambiguated and that this disambiguation generates predictable durational differences. A relatively faster tempo/shorter duration is found for less predictable (and harder to parse) nested structure, such as (reduced-)RCs, than the more predictable sisterhood structure, such as a main clause, CC, and pseudo-relative (in Italian). These results strongly suggest that the effects of both prosodic structure and predictability on duration are not always aligned. When not aligned, structural factors seem to determine durational properties above and beyond predictability. We have provided a principled account of these effects and argued that, while surprising at first sight, these results are expected to arise from the application of independently motivated principles of prosody.

We do not mean to suggest that the disambiguation of nested structure is achieved solely on the basis of durational information. Prosody varies along multiple dimensions (duration/tonal changes/intensity), and a combination of any or all of these can be and (as our preliminary results show) is used for encoding prosodic structure. More work is thus needed to fully establish the relative contribution of these different factors to the disambiguation of nested garden path sentences and to firmly establish to what extent durational differences are decisive. It's also still very much an open question whether the effects observed here are best explained from a localized perspective (in which the durational differences should be taken to reflect boundary phenomena) or whether global accounts (in which temporal differences are expected beyond boundary regions) should also be invoked. Which aspect of duration is more relevant – that across boundary lengthening or changes to speech rate – needs to be further explored.

We conclude by stressing once more that the global effects on speech rate are compatible with the independently motivated localized pattern of pre- and post-boundary lengthening described above, and in fact the two factors appear to show independent effects at different regions. While more work is needed to properly assess the hypothesis presented here and disentangle the relative contribution of these two factors, we have sketched a principled argument for garden paths with a temporary clausal attachment ambiguity to lead to a shorter duration for the less predictable nested structure.

Box 19.4 Chapter Overview

Summary

Tracking temporal properties of speech is essential to parsing both segmental and suprasegmental information, and changes in tempo can inform syntactic processing. Results from production and comprehension show that nested garden path sentences (e.g., *the horse raced past the barn (and) fell*) are prosodically disambiguated. This is achieved (also) through tempo modulation. Tempo modulation is a hallmark of structural nesting, generated by the interaction of eurhythmic and syntax–prosody mapping principles.

https://doi.org/10.1017/9781009295888.023 Published online by Cambridge University Press

Box 19.4 (cont.)

Implications

Nested garden path sentences are exceptional in displaying relatively faster tempo in the face of lower predictability. Future work should focus on the interplay of tempo and other prosodic variables in nesting and on how these are independently modulated by information structure, predictability, and constituent structure.

Gains

These findings have implications for cognitive science and psycholinguistics in countering standard simple views that higher predictability maps onto a shorter duration. These effects can be overshadowed by prosodic factors, as an increased tempo encodes the less predictable nested structure. The work is relevant to a growing body of literature on the role of rhythm in neurobiology of language by showing that changes in rhythm are a key signal from the producer to the comprehender of differential structure to the same linear word order.

REFERENCES

Altmann, G. T. M., and Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, 30, 191–281.

Aylett, M. P. (2000). Stochastic suprasegmentals: Relationships between redundancy, prosodic structure and care of articulation in spontaneous speech. PhD thesis, University of Edinburgh.

Aylett, M. P., and Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech*, 47, 31–56.

Beckman, M. E., and Ayers, G. (1997). Guidelines for ToBI labelling. *OSU Research Foundation*, 3, 255–309.

Beckman, M. E., and Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. *Phonology*, 3, 255–309.

Bennett, R., and Elfner, E. (2019). The syntax–prosody interface. *Annual Review of Linguistics*, 5, 151–171.

Bever, T. G. (1970). The cognitive basis of linguistic structure. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 279–362). New York: Wiley.

Breen, M., Watson, D. G., and Gibson, E. (2011). Intonational phrasing is constrained by meaning, not balance. *Language and Cognitive Processes*, 26, 1532–1562.

Carlson, K., Clifton, C. Jr., and Frazier, L. (2001). Prosodic boundaries in adjunct attachment. *Journal of Memory and Language*, 45, 58–81.

Cho, T., and Keating, P. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics*, 37, 466–485.

https://doi.org/10.1017/9781009295888.023 Published online by Cambridge University Press

Clifton, C. Jr., Carlson, K., and Frazier, L. (2002). Informative prosodic boundaries. *Language and Speech*, 45, 87–114.

Clifton, C. Jr., Traxler, M. J., Mohamed, M. T., et al. (2003). The use of thematic role information in parsing: Syntactic processing autonomy revisited. *Journal of Memory and Language*, 49, 317–334.

Cooper, W. E., and Paccia-Cooper, J. (1980). *Syntax and speech*. Cambridge, MA, and London: Harvard University Press.

Connine, C., Ferreira, F., Jones, C., Clifton, C. Jr., and Frazier, L. (1984). Verb frame preferences: Descriptive norms. *Journal of Psycholinguistic Research*, 13, 307–319.

Crain, S., and Steedman, M. (1985). On not being led up the garden path: The use of context by the psychological syntax processor. In D. Dowty, L. Karttunen, and A. Zwicky (Eds.), *Natural language parsing* (pp. 320–358). Cambridge, MA: Cambridge University Press.

Cutler, A. (1976). Phoneme-monitoring reaction time as a function of preceding intonation contour. *Perception & Psychophysics*, 20, 55–60.

Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*. Cambridge, MA: MIT Press.

Cutler, A., and Fodor, J. A. (1979). Semantic focus and sentence comprehension. *Cognition*, 7, 49–59.

Dahan, D. (2015). Prosody and language comprehension. *Wiley Interdisciplinary Reviews: Cognitive Science*, 6, 441–452.

Davis, M. H., Marslen-Wilson, W. D., and Gaskell, M. G. (2002). Leading up the lexical garden path: Segmentation and ambiguity in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 218.

Ferreira, F., and Clifton, C. Jr., (1986). The independence of syntactic processing. *Journal of Memory and Language*, 25, 348–368.

Fodor, J. D. (2002). Psycholinguistics cannot escape prosody. Proceedings of Speech Prosody 2002, pp. 83–90. Aix-en-Provence, France.

Fodor, J. D., Nickels, S., and Schott, E. (2018). Center-embedded sentences: What's pronounceable is comprehensible. In R. de Almeida, and L. Gleitman (Eds.), *On concepts, modules and language: Cognitive science at its core* (pp. 139–168). New York: Oxford University Press.

Frazier, L. (1979). On comprehending sentences: Syntactic parsing strategies. Doctoral dissertation, University of Connecticut.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 559–585). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Frazier, L., and Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, 6(4), 291–325. https://doi.org/10/cfc7p7

Frazier, L., Carlson, K., and Clifton, C. Jr., (2006). Prosodic phrasing is central to language comprehension. *Trends in Cognitive Sciences*, 10, 244–249.

Gee, J. P., and Grosjean, F. (1983). Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology*, 15, 411–458.

Ghini, M. (1993). Phi-formation in Italian: A new proposal. *Toronto Working Papers in Linguistics*, 12, 41–78.

Greenberg, S. (1999). Speaking in shorthand: A syllable-centric perspective for understanding pronunciation variation. *Speech Communication*, 29, 159–176.

Greenberg, S. (2006). A multi-tier framework for understanding spoken language. In S. Greenberg and W. A. Ainsworth (Eds.), *Listening to speech: An auditory perspective* (pp. 411–433). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Greenberg, S., Carvey, H., Hitchcock, L., and Chang, S. (2003). Temporal properties of spontaneous speech: A syllable-centric perspective. *Journal of Phonetics*, 31(3–4), 465–485. https://doi.org/10.1016/j.wocn.2003.09.005

https://doi.org/10.1017/9781009295888.023 Published online by Cambridge University Press

Grillo, N., and Turco, G. (2016). Prosodic disambiguation and attachment height. Proceedings of Speech Prosody 2016, pp. 1176–1180. https://doi.org/10.21437/SpeechProsody.2016-242

Grillo, N., Aguilar, M., Roberts, L., Santi, A., and Turco, G. (2018). Prosody of classic garden path sentences: The horse raced faster when embedded. Proceedings of Speech Prosody 2018, pp. 284–288. https://doi.org/10.21437/SpeechProsody.2018-58

Grillo, N., Aguilar, M., Roberts, L., Santi, A., and Turco, G. (2019). Less predictable and faster: Predictability and duration in Embedding and Sisterhood. Psycholinguistics in Iceland: Parsing and Prediction (PIPP) Conference, Reykjavík, Iceland, June 19–20.

Grillo, N., Aguilar, M., Roberts, L., Santi, A., and Turco, G. (2023). Garden path no more: Prosodic disambiguation of complement clauses and relative clauses. Poster, AMLaP 29 (Architectures and Mechanisms for Language Processing). Donostia, Spain.

Grillo, N., Santi, A., Aguilar, M., Roberts, L., and Turco, G. (2022). Prosodic phrasing leads to shorter duration for more complex structures: The case of garden path sentences. https://ssrn.com/abstract = 4008098

Guo, B., Grillo, N., Mattys, S., et al. (2023). Prosody disambiguates string-identical connected clauses and relative clauses. Poster, AMLaP 29 (Architectures and Mechanisms for Language Processing). Donostia, Spain.

Guo, B., Grillo, N., Mattys, S., et al. (2024a). The prosody of clefted relatives: A new window into prosodic representations. Proceedings of Speech Prosody 2024, Leiden, the Netherlands, pp. 1215–1219. https://doi.org/10.21437/SpeechProsody.2024-245

Guo, B., Grillo, N., Mattys, S., et al. (2024b). Phrasing and prominence disambiguate clefted Relative Clauses. Under review in *Laboratory Phonology*.

Guo, B., Santi, A., Turco, G. and Grillo, N. (in preparation). Rhythmic changes modulate the strength of garden-path effects. Ms. University of York.

Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. Proceedings of the Second North American Chapter of the Association for Computational Linguistics on Language Technologies (NAACL '01). Association for Computational Linguistics, USA, pp. 1–8. https://doi.org/10.3115/1073336.1073357

Harris, J., Jun, S.-A., and Royer, A. (2016). Implicit prosody pulls its weight: Recovery from garden path sentences. Proceedings of Speech Prosody 2016, pp. 207–211. https://doi.org/10.21437/SpeechProsody.2016-43

Hirschberg, J., and Avesani, C. (1997). The role of prosody in disambiguating potentially ambiguous utterances in English and Italian. Proceedings of Intonation: Theory, Models and Applications, pp. 189–192.

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, 20, 137–194.

Jurafsky, D., Bell, A., Gregory, M., and Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. *Typological Studies in Language*, 45, 229–254.

Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition*, 2, 15–47.

Kjelgaard, M. M., and Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language*, 40, 153–194.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, 59, 1208–1221.

Konieczny, L. (2000). Locality and parsing complexity. *Journal of Psycholinguistic Research*, 29, 627–645. https://doi.org/10.1023/A:1026528912821

Kraljic, T., and Brennan, S. E. (2005). Prosodic disambiguation of syntactic structure: For the speaker or for the addressee? *Cognitive Psychology*, 50, 194–231.

https://doi.org/10.1017/9781009295888.023 Published online by Cambridge University Press

Kuperberg, G. R., and Jaeger, T. F. (2015). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, 31(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299

Ladd, D. R. (1986). Intonational phrasing: The case for recursive prosodic structure. *Phonology*, 3, 311–340.

Lau, E., Stroud, C., Plesch, S., and Phillips, C. (2006). The role of structural prediction in rapid syntactic analysis. *Brain and Language*, 98(1), 74–88. https://doi.org/10.1016/j.bandl.2006.02.003

Lehiste, I. (1972). The timing of utterances and linguistic boundaries. *Journal of the Acoustical Society of America*, 51, 2018–2024.

Levy, R. P. (2008). Expectation-based syntactic comprehension. *Cognition*, 106(3), 1126–1177. https://doi.org/10/dtkt7j

Levy, R. P. (2011). Integrating surprisal and uncertain-input models in online sentence comprehension: Formal techniques and empirical results. In D. Lin, Y. Matsumoto, and R. Mihalcea (Eds.), *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies* (pp. 1055–1065). Portland, OR: Association for Computational Linguistics.

Levy, R. P., and Jaeger, T. F. (2007). Speakers optimize information density through syntactic reduction. *Advances in Neural Information Processing Systems*, 19, 849.

Levy, R. P., Fedorenko, E., Breen, M., and Gibson, E. (2012). The processing of extraposed structures in English. *Cognition*, 122(1), 12–36. https://doi.org/10.1016/j.cognition.2011.07.012

Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech*, 6, 172–187.

MacDonald, M. C., Pearlmutter, N. J., and Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101(4), 676–703. https://doi.org/10.1037/0033-295X.101.4.676

McRae, K., Spivey-Knowlton, M. J., and Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, 38, 283–312.

Miller, J. L., and Grosjean, F. (1981). How the components of speaking rate influence perception of phonetic segments. *Journal of Experimental Psychology: Human Perception and Performance*, 7(1), 208–215.

Nespor, M., and Vogel, I. (2007). *Prosodic phonology: With a new foreword*. Berlin and Boston: De Gruyter Mouton.

Ni, W., Crain, S., and Shankweiler, D. (1996). Sidestepping garden paths: Assessing the contributions of syntax, semantics and plausibility in resolving ambiguities. *Language and Cognitive Processes*, 11, 283–334.

Nooteboom, S. (1997). The prosody of speech: melody and rhythm. *The Handbook of Phonetic Sciences*, 5, 640–673.

Paterson, K. B., Liversedge IV, S. P., and Underwood, G. (1999). The influence of focus operators on syntactic processing of short relative clause sentences. *Quarterly Journal of Experimental Psychology: Section A*, 52, 717–737.

Pickering, M. J., and Van Gompel, R. P. (2006). Syntactic parsing. In M. J. Traxler and M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (pp. 455–503). Amsterdam: Elsevier.

Pierrehumbert, J. B. (1980). The phonology and phonetics of English intonation. PhD thesis, Massachusetts Institute of Technology.

Poschmann, C., and Wagner, M. (2015). Relative clause extraposition and prosody in German. *Natural Language & Linguistic Theory*, 34, 1021–1066. https://doi.org/10.1007/s11049-015-9314-8

Reinisch, E., and Sjerps, M. J. (2013). The uptake of spectral and temporal cues in vowel perception is rapidly influenced by context. *Journal of Phonetics*, 41(2), 101–116.

https://doi.org/10.1017/9781009295888.023 Published online by Cambridge University Press

Reinisch, E., Jesse, A., and McQueen, J. M. (2011). Speaking rate from proximal and distal contexts is used during word segmentation. *Journal of Experimental Psychology: Human Perception and Performance*, 37(3), 978–996.

Salverda, A. P., Dahan, D., and McQueen, J. M. (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition*, 90, 51–89.

Santi, A., Grillo, N., Molimpakis, E., and Wagner, M. (2019). Processing relative clauses across comprehension and production: Similarities and differences. *Language, Cognition and Neuroscience*, 34(2), 170–189. https://doi.org/10.1080/23273798.2018.1513539

Schafer, A. J. (1997). Prosodic parsing: The role of prosody in sentence comprehension. PhD thesis, University of Massachusetts Amherst.

Schafer, A. J., Speer, S. R., Warren, P., and White, S. D. (2005). Prosodic influences on the production and comprehension of syntactic ambiguity in a game-based conversation task. In J. C. Trueswell and M. K. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions* (pp. 209–225). Cambridge, MA: MIT Press.

Sedivy, J. C. (2002). Invoking discourse-based contrast sets and resolving syntactic ambiguities. *Journal of Memory and Language*, 46, 341–370.

Selkirk, E. (1980). *On prosodic structure and its relation to syntactic structure, volume 194*. Bloomington, IN: Indiana University Linguistics Club.

Selkirk, E. (2000). The interaction of constraints on prosodic phrasing. In M. Horne (Ed.), *Prosody: Theory and experiment: Studies presented to Gösta Bruce* (pp. 231–261). Dordrecht: Springer.

Shatzman, K. B., and McQueen, J. M. (2006). Segment duration as a cue to word boundaries in spoken-word recognition. *Perception & Psychophysics*, 68, 1–16.

Silverman, K. E., Beckman, M. E., Pitrelli, J. F., et al. (1992). Tobi: A standard for labelling English prosody. ICSLP, pp. 867–870. Volume 2.

Snedeker, J., and Trueswell, J. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language*, 48, 103–130.

Speer, S., and Blodgett, A. (2006). Prosody. In M. J. Traxler and M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (pp. 505–537). Amsterdam: Elsevier.

Speer, S., Warren, P., and Schafer, A. J. (2011). Situationally independent prosodic phrasing. *Laboratory Phonology*, 2, 35–98.

Staub, A., and Clifton Jr, C. (2006). Syntactic prediction in language comprehension: Evidence from either... or. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(2), 425.

Staub, A., Foppolo, F., Donati, C., and Cecchetto, C. (2018). Relative clause avoidance: Evidence for a structural parsing principle. *Journal of Memory and Language*, 98, 26–44.

Tanenhaus, M. K., Carlson, G., and Trueswell, J. C. (1989). The role of thematic structures in interpretation and parsing. *Language and Cognitive Processes*, 4(3–4), SI211–SI234.

Traxler, M. J. (2014). Trends in syntactic parsing: Anticipation, Bayesian estimation, and good-enough parsing. *Trends in Cognitive Sciences*, 18(11), 605–611. https://doi.org/10.1016/j.tics.2014.08.001

Truckenbrodt, H. (1995). Phonological phrases: Their relation to syntax, focus, and prominence. PhD thesis, Massachusetts Institute of Technology.

Trueswell, J. C., Tanenhaus, M. K., and Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, 33, 285.

Turk, A., and Shattuck-Hufnagel, S. (2014). Timing in talking: What is it used for, and how is it controlled? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369, 20130395.

Wagner, M. (2005). Prosody and recursion. PhD thesis, MIT.

Wagner, M. (2010). Prosody and recursion in coordinate structures and beyond. *Natural Language and Linguistic Theory*, 28, 183–237.

https://doi.org/10.1017/9781009295888.023 Published online by Cambridge University Press

Wagner, M., and Watson, D. G. (2010). Experimental and theoretical advances in prosody: A review. *Language and Cognitive Processes*, 25, 905–945.

Watson, D., and Gibson, E. (2005). Intonational phrasing and constituency in language production and comprehension. *Studia Linguistica*, 59, 279–300.

Watson, D. G., Arnold, J. E., and Tanenhaus, M. K. (2008). Tic tac toe: Effects of predictability and importance on acoustic prominence in language production. *Cognition*, 106, 1548–1557.

White, L., and Turk, A. E. (2010). English words on the procrustean bed: Polysyllabic shortening reconsidered. *Journal of Phonetics*, 38, 459–471.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., and Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America*, 91, 1707–1717.

https://doi.org/10.1017/9781009295888.023 Published online by Cambridge University Press