Contents

This section goes over evidence supporting the claim that prosody can impact parsing decisions, even in silent reading.

1   Can prosody affect parsing?

A number of studies have shown that in listening to speech, prosodic cues appear to help reduce the frequency with which incorrect parsing (i.e. a garden path) occurs. For example, Kjelgaard and Speer (1999) conducted a study using digital manipulation of recorded speech to create three versions of sentences containing a garden path temporary ambiguity (discussed above). They recorded speakers saying sentences with natural prosody, such as the following pair:

(1) [When Roger leaves] the house is dark. (Early closure)
(2) [When Roger leaves the house] it's dark. (Late closure)

They then cross-spliced these together to make several versions. One version had prosodic cues which cooperated with the intended reading of the sentence; another attempted to have "neutral" prosody; and the third used intentionally misleading prosody. The initial fragment of each was then presented to participants (the portion from the beginning of the sentence to the word house in (1 - 2) and they were asked to agree or disagree with whether a visually presented word, either *is* or *it's* was likely to be the next word in the sentence. Participants gave more accurate and speedier judgements when the prosodic cues lined up with the correct parsing. The results of this study, as well as a growing body of literature, suggest that that prosodic information can (or perhaps must) be used by the parser in making processing decisions.

2   Prosody in silent reading

The experiment presented in this paper is conducted using silent reading, and it is perhaps less intuitive to assume that prosody might play a role in the processing of silent reading compared to the processing of spoken utterances. Nonetheless, there is some evidence to support such an effect; as such, a positive result here is not a reason to dismiss the prosodic explanation for the effect, although it also can't support that explanation over others. Consider Fodor (2002) analysis of relative clause attachment preference. This concerns sentences such as (3):

(3) Someone shot the servant$_{N1}$ of the actress$_{N2}$ [$_{RC}$ who was on the balcony].

The relative clause (RC) who was on the balcony can attach either locally (low, modifying N2), making it *the actress* who was *on the balcony*, or higher up (non-locally, modifying N1), so that we understand *the servant* to be the one who was *on the balcony*. In these sorts of sentences, Cuetos and Mitchell (1988) found a 60% preference for low attachment in English speakers, but only a 40% preference for low attachment in Spanish speakers. In apparent violation of the general preference for local attachment, recall *Minimal Attachment* (??), from Frazier (1979), some languages, like French and Spanish (and Russian, but not Romanian or Brazilian Portuguese, so this is not a general feature of Romance languages), prefer to attach relative clauses higher, while others more often obey Late Closure (e.g., Swedish, Egyptian Arabic, and English). This non-local preference is weakened in cases where the ambiguous RC is short (one prosodic word). Fodor (2002) asserts that these tendencies exist in both listening to spoken words (under conditions where a particular parse is not favored by the explicit prosody) and in silent reading.

Fodor notes that other researchers have shown the presence and absence of prosodic breaks to influence parsing decisions, and specifically that the presence of a prosodic break before the RC in sentences like (5) encourages high attachment. Fodor leverages this in order to explain the difference in RC attachment site tendency between languages. She argues that the phenomenon can be neatly account for by linking attachment site preference to the likelihood of a prosodic break before the RC. This difference in prosodic tendency, in turn, can be explained using a constraints-based approach. Consider Selkirk's (1986) alignment constraints:

Definition 2.1 (Generalised Alignment)**.** Align(αCat, E; βCat, E)

    a. Align (GCat, E; PCat, E)
    b. Align (PCat, E; GCat, E)
    c. Align (PCat, E; PCat, E)

GCat ranges over morphological and syntactic categories; PCat ranges over prosodic categories; E = Right or Left (Selkirk, 1986, p. 6)

Truckenbrodt (1999) provides a prose-based formalization of the same idea. He describes what I will call $Align_R$ which can be easily generalized to described what I will call $Align_L$, the same constraint except that it calls for aligning phrases at their left edges rather than their right edges.

  (4) **Align-XP/R** For each XP there is a PP such that the right edge of the XP coincides with the right edge of the PP, where XP is a maximal projection and PP is a Phonological Phrase. This constraint represents the end based mapping assumption for Major Phonological Phrases in English, whose right end is supposed to align with the right end of Maximal Projections (Truckenbrodt, 1999, p. 223).

Essentially, Selkirk argues that relative ranking of alignment constraints for the left edge of phrases ($Align_L$) with those for the right edge of phrases ($Align_R$) can impact the distribution of prosodic breaks. These alignment constraints dictate that the edges of prosodic units (and thus the location of prosodic breaks) should align with the edges of syntactic constituents. Because the prosodic break that encourages high attachment is one which aligns with the left edge of the RC, postulating that $Align_L$ is ranked above $Align_R$ in languages like French that prefer high attachment can account for that preference (remember that a prosodic break in that place has been shown to encourage a high attachment interpretation). In languages where low attachment is preferred, we can assume that $Align_R$ is ranked higher, and thus a prosodic break is more likely to occur after the RC than before.

The same sort of argument can explain the difference in tendency between long and short RCs. Consider Selkirk's (2011) *BinMin* defined below.

  (5) **BinMin(ϕ)** A ϕ (phonological phrase) must consist of at least two ω (phonological words).

If we assume, in Optimality Theoretic (Prince & Smolensky, 1993) terms, that a constraint like *BinMin* is ranked above $Align_L$, then it seems quite reasonable to assume that a prosodic break before a short RC (which would encourage high attachment) is much less likely than before a long RC. That is, when the RC is short, its left edge is prevented from aligning with the beginning of a prosodic phrase (it violates $Align_L$) by the higher ranked BinMin. Longer RCs can have their left edge align with the start of a prosodic phrase, and thus can have the high-attachment encouraging prosodic break.

This analysis, of course, assumes the presence of prosodic cues, which are only overtly present in spoken utterances (punctuation aside). Given the elegance and explanatory power of this analysis, it seems reasonable to at least consider the further assumption necessary for it to also explain the phenomenon as it occurs in silent reading; which is that readers project prosodic structure over sentences even when they are read silently.

3   Prosody of questions vs. declaratives

In pursuing the possibility that it is the intonation and prosody of polar interrogatives which creates the intuitive contrast that this study investigates, we must consider what question intonation actually sounds like.

It is generally agreed that in American English, the intonation of a polar (yes/no) question has the property of a final rise. Indeed, this has been confirmed in corpus studies such as Hedberg et al. (2017) who found that 79.8% of the 410 American English yes/no questions in their study (ten-minute phone conversations from the CallHome Corpus of American English and the Fisher English Corpus) had a "low-rise nuclear contour" (L*H-H%, L*H-↑H%, or L*L-H%)[1]. To briefly explain their ToBI notation, a tone T is either L for low or H for high; T* is anchored to the stressed syllable, and T- and T% are boundary tones (intermediate phrase boundary and intonational phrase boundary respectively). See, e.g., *Guidleines for ToBI labeling* (Beckman and Ayers, 1997) for a more thorough explanation of ToBI. An additional 10.7% of the Hedberg et al. (2017) data had a "high-rise nuclear contour" (the authors categorizes the following tunes as "high-rise nuclear contours:" H*H-H%, or !H*L-L%). That leaves only 9.5% spread across the other 5 categories (High-fall, Rise-fall, Low-fall, Fall-rise, and Level). Only 5.6% of the data showed a falling contour. According to the authors' analyses, these contours occur on the final main stress of a sentence and thereafter. In the case of the types of sentences examined in the current study,that would result in a rising contour on the head noun of the final PP.

(6) Did Jed cram the newspapers under the sofa in the [$_{\text{L*H-H\%}}$ guestroom].

The need to prepare for that rising tone might make a prosodic break before the PP more likely, and thus ease reanalysis or even encourage a different prosodic chunking which might encourage high attachment.

A brief informal survey found that most speakers maintain low tones on prior stresses, although some had a H tone on the subject noun. It also varied between speakers and between sentences as to whether there is a prosodic boundary (marked by a low tone and/or pause) immediately before the rise (after PP1) or not.

The prosodic structures found in the data collected for the current study are discussed in **??**.

4   Syntactic structure of PP-attachment ambiguities

As discussed in the introduction, this study is focused on the impact of sentence type (interrogative vs. declarative) in a particular sort of garden path. Specifically, it is concerned with garden path sentences containing a temporary ambiguity that centers on the attachment of prepositional phrases (PPs) occurring in string-linear sequence at the end of an utterance, e.g., "When he saw her, the nanny had seated the cranky little boy [PP1 on the swing] [PP2 in his stroller]." Consider (13):

(7) Jed crammed the newspapers under the sofa in the trashcan.

    a. # … [$_{\text{VP}}$ crammed [$_{\text{NP}}$ the newspapers] [$_{\text{PP1}}$ under [$_{\text{NP}}$ the sofa [$_{\text{PP2}}$ in the trashcan]]]
    b. ✓… [$_{\text{VP}}$ crammed [$_{\text{NP}}$ the newspapers [$_{\text{PP1}}$ under [$_{\text{NP}}$ the sofa]] [$_{\text{PP2}}$ in the trashcan]]

                                                   # indicates a structure with an implausible reading

In the case of this sort of garden path, Minimal Attachment, or some variation thereof, results in a structure where the first PP attaches to the verb phrase (VP) as an argument, i.e. [$_{\text{VP}}$ V NP PP1], which leaves nowhere for the second PP to attach but as a modifier of the noun phrase (NP) inside PP1 (#[$_{\text{PP1}}$ on [$_{\text{NP}}$ the swing [$_{\text{PP2}}$ in his stroller]]]). This initial parse (7~ ~a) is pragmatically implausible, as one does not generally find swings inside strollers. Reanalysis is required to bring about the correct parse (7~ ~b), where PP1 attaches as an NP modifier of the direct object [$_{\text{VP}}$ V [$_{\text{NP}}$ N PP1] …] which allows PP2 to attach as a VP argument, i.e. it is the cranky little boy on the swing who is being placed in the stroller.

Note that Minimal Attachment as defined by Frazier (1979) is somewhat at odds with modern syntactic theory (cf. Chomsky 1995, p. 62). As originally postulated, it relies on a verb with multiple internal arguments incorporating each of those arguments as a sister (i.e. a ternary branching structure: [VP V NP PP]). With more modern theories where binary branching is obligatory, it becomes less clear that the VP attachment site for PP1 actually creates fewer nodes than the lower NP attachment site. Nonetheless, the preference for VP attachment in these kinds of sentences is there, be it due to Minimal Attachment, a preference for arguments over non-arguments, or whatever else: see evidence from e.g. Rayner et al. (1983) and Clifton et

---

[1]Hedberg et al. (2017) use ↑ to indicate an upstep, which is not standardly transcribed with ToBI.

al. (1991). I assume, following Frazier (1998) that it can be adequately described by reformulating Minimal Attachment to include a preference for argument relations over non-argument ones, or else through careful consideration of the syntactic structures involved and reformulation of Minimal Attachment. The precise definition of this parsing strategy is not a concern of this paper, but it is assumed to be at play.

References

Beckman, M. E. and Ayers, G. (1997). Guidelines for ToBI labelling. *The OSU Research Foundation*, 3:30.

Cuetos, F. and Mitchell, D. C. (1988). Cross-linguistic differences in parsing: restrictions on the use of the late closure strategy in spanish. *Cognition*, 30(1):73–105.

Fodor, J. D. (2002). Psycholinguistics cannot escape prosody. In *Speech Prosody 2002, International Conference*.

Frazier, L. (1979). *On Comprehending Sentences: Syntactic Parsing Strategies*.

Hedberg, N., Sosa, J. M., and Görgülü, E. (2017). The meaning of intonation in yes-no questions in american english: A corpus study.

Kjelgaard, M. M. and Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity.

Selkirk, E. O. (1986). *Phonology and Syntax: The Relation Between Sound and Structure*. MIT Press (MA).

Selkirk, E. O. (2011). The syntax-phonology interface. In Goldsmith, J., Riggle, J., and Yu, A., editors, *The Handbook of Phonological Theory*, volume 2, pages 435–483. Oxford: Blackwell Publishing.

Truckenbrodt, H. (1999). On the relation between syntactic phrases and phonological phrases.