

Prepositional phrase attachment
ambiguities in declarative and
interrogative contexts: Oral reading
data

Tyler J. Peckenpaugh

2019-08-17

Contents

Abstract	vii
Acknowledgements	viii
About this draft	viii
1 Introduction and background	1
1.1 Motivations for the current	4
2 Background	7
2.1 Structural overview of the ambiguity relevant to this study	8
2.2 Interrogativity	16
2.3 Prosody of questions vs. declaratives	19
2.4 Can prosody affect parsing?	20
2.5 Predictions for the current study	24
3 Methodology	27
3.1 Materials	28
3.2 Participants recruitment	35
3.3 Location	35
3.4 Equipment and software	36
3.5 Versions of the experiment	36
3.6 Procedure	37

3.7	Measurements of utterance timing	44
3.8	Prosodic judgments	47
4	Results and discussion	50
4.1	Data for analysis	51
4.2	Prosodic break patterns	52
4.3	Discussion of prosodic break patterns	64
4.4	Inter-reading time	68
4.5	Discussion of IRT results	71
5	General discussion	75
5.1	Behavioral correlate for the intuition	76
5.2	On explanations for the intuition	76
5.3	Conclusions	79
5.4	Summary	83
A	Experimental items	85
B	Filler items	88
C	Recruitment notice	92
D	Instructions to participants	93
E	Instructions for prosodic coding	95
	References	97

List of Tables

3.1	Illustrative experimental item, constructed in four versions.	28
3.2	Illustrative filler items, constructed in two versions.	34
3.3	Distance in characters from fixation to disambiguation in experimental items.	42
3.4	EVS-adjusted character distance to disambiguation in experimental items.	43
3.5	Inter and intra-rater agreement.	49
4.1	Number of participants per version-order combination.	51
4.2	Number of recordings analyzed, as a function of Speech Act and PP2 Status.	52
4.3	Percent occurrence of OBJ break (frequency of occurrence in parenthesis) as a function of sentence type and Reading.	53
4.4	Percent occurrence of PP1 break (frequency of occurrence in parenthesis) as a function of sentence type and Reading.	53
4.5	Percent occurrence of both breaks as a function of sentence type and Reading.	54
4.6	Mixed effects logistic regression model predicting OBJ break occurrence (FULL).	57

4.7	Mixed effects logistic regression model predicting OBJ break occurrence (REDUCED).	58
4.8	Mixed effects logistic regression model predicting PP1 break occurrence (FULL).	58
4.9	Mixed effects logistic regression model predicting PP1 break occurrence (REDUCED).	59
4.10	Mixed effects logistic regression model predicting OBJ break dominance (FULL).	60
4.11	Mixed effects logistic regression model predicting OBJ break dominance (REDUCED).	60
4.12	Mixed effects logistic regression model predicting PP1 break dominance (REDUCED).	60
4.13	Mixed effects logistic regression models predicting break dominance in Reading 2 (REDUCED).	61
4.14	Break pattern by sentence type and R1 delay category.	64
4.15	Occurrence of both breaks as a function of Reading, PP2 Status, and R1 Delay Category.	64
4.16	Linear mixed effects regression model predicting wIRT by sentence type with interaction term (FULL).	70
4.17	Linear mixed effects regression model predicting wIRT by sentence type (REDUCED).	70
4.18	Linear mixed effects regression model predicting wIRT by Speech Act and PP2 head.	71
5.1	Percent occurrence of simple break patterns in Reading 2 as a function of sentence type.	77
A.1	Experimental items in four versions	85

B.1	Filler items with trailing PPs	88
B.2	Filler items with no trailing PPs	91
D.1	Table of keyboard mappings	94

List of Figures

2.1	Syntactic tree of an illustrative example sentence with an ambiguous PP1 and a modifier-PP2.	12
2.2	Syntactic tree of an illustrative example sentence with an ambiguous PP1 and an argument-PP2.	13
3.1	Diagram of 4-screen sequence presented for each item, showing the key presses triggering movement between successive screens.	38
4.1	Break pattern as a function of sentence type and Reading.	54
4.2	Percent break dominance occurrence as a function of sentence type and Reading.	55
4.3	Distributions of R1 delay and R2 delay	62
4.4	Plot of pattern proportions as a function of sentence type.	63
4.5	Distribution of wIRT.	69
4.6	Mean IRT as a function of sentence type.	69

Abstract

Abstract is a work in progress. This paper reports a study on the effect of interrogativity on the oral reading of temporarily ambiguous prepositional phrases (PPs). Specifically, it looks at sentences ending in a of two PPs, where the first is interpretable as the goal argument of the preceding verb, and the status of the second (PP2 Status) is manipulated to either necessarily be the goal argument of that verb (Arg), forcing reanalysis, or not (Mod), allowing the original parse to stand. No evidence is found that interrogativity impacts the difficulty of understanding the Arg-type sentences, despite an intuitive decrease in difficulty when those sentences are presented in an interrogative context. A double-reading protocol is employed, where participants are asked to read a sentence first without preview (Reading 1), and then after unlimited preview (Reading 2). A robust effect of PP2 Status is found for the prosodic phrasing of the target sentences, and an effect of interrogativity on the study time between Readings, Inter-Reading Time (IRT), is reported.

Acknowledgements

TBD

About this draft

This represents the document that will be defended on August 29th. The goal is to deposit before September 16, after incorporating whatever revisions are requested.

Chapter 1

Introduction and background

This section represents a combination of what was the start of chapter 1 (introduction) and the start of chapter 2 (background). It is mostly text you've read, but substantially reorganized, and should be skimmed and discussed.

This paper presents a study on human sentence processing, or parsing, and on the parsing of a particular sort of ambiguity. Parsing is assumed to be the projection of structure by a reader or listener over a string (which obviously lacks inherent structure). Following the sort of model of parsing put forward by, e.g., Kimball (1973), Frazier & Fodor (1978), and Frazier & Clifton (1996), this study assumes that parsing is done online and that most material must be incorporated into the structure being built as soon as it is encountered. This can lead to mis-parses, where the parser has guessed wrong about how to incorporate a given phrase with a temporarily ambiguous structure, and then encounters material that cannot be incorporated into the current structure. When this happens, the parser must reanalyze the material that had so far been processed, in order to come up a structure that can accomodate both the new and old material grammatically.

[1]: This paragraph was ¶1 chapter 1

part or all of its pending structure and try again to incorporate the material encountered so far. This sort of parser crash is often called a garden path.

As briefly mentioned earlier, “Garden path effects” occur when a temporarily ambiguous sentence resolves in such a way that the structure initially preferred by the parser is incompatible with how the sentence actually continues. These parsing errors have traditionally been attributed to structurally-focused parsing preferences (Frazier, 1979; Frazier & Fodor, 1978; Kimball, 1973) that ignore semantic content on the first pass. Frazier (1979) formulates several of these, including the following two which are widely accepted in one form or another:

[2]: From here until next note, was introductory section of chapter 2

- (1) *Minimal attachment* Attach incoming material into the phrase-marker being constructed using the fewest nodes consistent with the well-formedness rules of the language under analysis (Frazier, 1979, p. 24)
- (2) *Late closure* When possible, attach incoming material into the clause currently being parsed (Frazier, 1979, p. 20)

Because these strategies ignore semantic and pragmatic plausibility and the parser typically does not know what material might be further on in the string, mis-parses at temporarily ambiguous regions can occur, resulting in garden paths. *Minimal Attachment* is important to this study and will be revisited later on.

An example is the commonly studied garden path sentence, “The horse raced past the barn fell” (Bever, 1970). Here, the initial parse incorrectly assumes that the matrix subject is the unmodified NP the horse, per Minimal Attachment, and takes the matrix verb to be raced, as in the sentence, “The horse raced past the finish line.”

- (3) The horse raced past the barn fell (Bever, 1970)

- a) [_S [_{NP} The horse] [_{VP} raced past the barn]] ??? [_{VP} fell]
- b) [_S [_{NP} The horse raced past the barn] [_{VP} fell]]

An attempted parse resulting in structure (41 a) crashes, as it is not possible to incorporate the final word *fell* in a grammatical way. Reanalysis is required, with the grammatical parse being (41 b) where the matrix subject is *the horse raced past the barn*, a noun phrase (NP) containing a reduced relative clause *raced past the barn*. Thus *fell* can be incorporated as the matrix verb, with a structure comparable to, “The horse (that was) raced past the barn was hungry.”

~~There is an ongoing debate in the literature about what parsing model best fits the empirical facts. This study follows [Frazier 1996] in assuming that structure-first parsing strategies are at play, in addition to a primary vs. non-primary relation distinction that determines how immediately a phrase must be incorporated into a parse, allowing for some material to be incorporated later and thereby make use of additional information that is not available for immediate parsing decisions.~~

The study being reported is concerned with certain sentences that contain *such* a temporarily ambiguous prepositional phrase (PP1), followed by another (PP2) which ~~causes the expected parse to crash~~ *interferes with what is assumed to be the default interpretation of PP1*. Specifically, it is expected that PP1 in (4) will initially be interpreted as the goal of *cram*, *but that parse will fail when it is realized that PP2 cannot plausibly modify drawer*.

[3]: Back to what was chapter 1, ¶ 2

(4) He had planned to cram the paperwork [PP1 in the drawer] [PP2 into his briefcase].

[4]: examples moved up slightly

(5) He had planned to cram the paperwork [PP1 in the drawer] [PP2 of his filing cabinet]].

This contrasts with the similar sentence in (5) where PP2 can plausibly modify *drawer* and so the parse where *in the drawer* is the goal argument is accepted and *of his filing cabinet* is incorporated as a modifier within PP1.} { as an argument of

the verb. The reasons for that assumption are discussed in Section 2.1, but for now it suffices to understand that (4) represents a difficult to comprehend sentence (a garden path), while (5) presents no such difficulty.}

((examples were here))

~~Before the details of the current research can be outlined, it is first necessary to explain some of the terms and mechanisms involved. This chapter is concerned with doing so.~~

\begin{note} Sections 1.1 and 1.2 are redundant and have been removed. They are incorporated into the above. \end

1.1 Motivations for the current

The current study was initially motivated by an observation discovered by Janet Dean Fodor and Dianne Bradley, and originally reported in @qp2. That observation is that a garden path sentence like (4) repeated here as (6) is, for whatever reason, not as difficult to process when presented as an interrogative, as in (7), rather than a declarative. ~~These attachment ambiguities which are difficult to parse in the declarative, appear to be less difficult to parse when presented in polar interrogative context. Consider again the somewhat difficult to process sentence in (@cramiGPy), repeated below as (@dec) for convenience.~~

[5]: Adapted from what was §1.3

(6) He had planned to cram the paperwork [_{pp1} in the drawer] [_{pp2} into his
briefcase].

(7) Had he planned to cram the paperwork [_{pp1} in the drawer] [_{pp2} into his
briefcase]?

Peckenpough (2016) attempted to find a behavioral correlate of this intuition by looking at variation in reading time of sentence like (6) and (7), but the results were