



The Prosody of Algebra and the Algebra of Prosody: Prosodic Disambiguation of Read Mathematical Formulae

Michael Phelan¹

¹Department of Linguistics, Ohio State University, USA

phelan@ling.osu.edu

Abstract

Read mathematical formulae (MF) provide an ideal and little-studied window into speakers' ability to prosodically disambiguate complex NPs. An experiment investigating whether and how speakers of different mathematical skill levels prosodically mark English utterances containing ambiguities caused by the application of unary and binary mathematical operators found that speakers consistently mark intended syntactic parses with *cooperating prosody*, whereby placement of relatively weaker prosodic breaks indicated tighter syntactic grouping and stronger breaks indicated looser syntactic grouping. When prosodic breaks were of equal strength, absolute break strength often indicated the desired parse. Use of cooperating prosody correlated significantly with correctly evaluating MF, and increased mathematical ability correlated with decreased use of conflicting (non-cooperating) prosodic contours.

Index Terms: speech prosody, relative boundary hypothesis, break strength, prosody of mathematics

1. Introduction

Mathematics is said to be a language unto itself, with a specialized vocabulary of functions, variables, and operators, and syntactic structures that can be deeply, even infinitely, recursive. This study extends the metaphor to investigate the prosody of read mathematical formulae (MF), and uses what is learned to argue for a particular theory of prosodic processing as an aid to syntactic ambiguity resolution.

MF possess a number of qualities that make them ideal stimuli for researchers interested in prosodic disambiguation of syntactic ambiguities. First, the syntactic and semantic relationships between terms in an MF are completely unambiguous given the written form of the MF, but it is easy to choose MF so that the corresponding segmental string used to read the MF contains ambiguities, as in (1).

- (1) a. Nine times A minus two: $(9 \cdot A) - 2$ $9 \cdot (A - 2)$
 b. Seven plus A squared: $(7 + A)^2$ $(7 + A^2)$
 c. The square root of A plus five: $\sqrt{A+5}$ $\sqrt{A+5}$

Second, unlike more commonly used stimuli, no particular parse of the segmental string corresponding to an MF is inherently more or less plausible than any other. Different speakers in the same speech community may also have very different levels of mathematical fluency. While it would be exceedingly rare to find adults who are not implicitly familiar with syntactic ambiguities caused by prepositional phrase or relative clause attachment, many speakers have less exposure to ambiguities in MF. This means that by using MF we can easily investigate the effect that experience with a certain type of ambiguity has on prosodic productions of sentences containing that ambiguity. Finally, the semantics of MF are cross-linguistically stable, and syntactic structures may be more likely to remain unchanged across languages.

2. Background and Assumptions

2.1. Defining the ambiguity

The ambiguities of interest in the present study are all based on the syntactic structure of complex NPs, and are exemplified in (1). The mathematical operators used in this study are assumed to be syntactically similar to conjunctions with sortal restrictions forcing their arguments to be numbers, variables, or complex NPs constructed of numbers and variables. They are assumed to have the following syntactic types:

- (2) a. (NP \ NP) / NP: plus, minus, times, divided by
 b. NP \ NP: squared
 c. NP / NP: the square root of...

Given this assumption, the ambiguities in the English segmental strings in (1) come about due to differences in the order in which the two operators in each string are combined with their operands, and are represented as:

- (3) a. $(9 \cdot A) - 2$ $[[\text{Nine}]_{\text{NP}} \text{ times } [A]_{\text{NP}}]_{\text{NP}} \text{ minus } [\text{two}]_{\text{NP}}$
 a'. $9 \cdot (A - 2)$ $[\text{Nine}]_{\text{NP}} \text{ times } [[A]_{\text{NP}} \text{ minus } [\text{two}]_{\text{NP}}]_{\text{NP}}$
 b. $(7 + A)^2$ $[[\text{Seven}]_{\text{NP}} \text{ plus } [A]_{\text{NP}}]_{\text{NP}} \text{ squared}$
 b'. $(7 + A^2)$ $[\text{Seven}]_{\text{NP}} \text{ plus } [[A]_{\text{NP}} \text{ squared}]_{\text{NP}}$
 c. $\sqrt{A+5}$ $[\text{The square root of } [A]_{\text{NP}}]_{\text{NP}} \text{ plus } [\text{five}]_{\text{NP}}$
 c'. $\sqrt{A+5}$ $\text{The square root of } [[A]_{\text{NP}} \text{ plus } [\text{five}]_{\text{NP}}]_{\text{NP}}$

In (3a,b,c) the first operator (*times*, *plus*, and *square root*, respectively) are applied as soon as the required number of operands are present. This is referred to as the *left-branching* [LB] *syntactic structure*. By contrast, in (3a',b',c'), the first operator is not applied immediately; the first operator instead combines with the complex NP resulting from applying the second operator (*minus*, *squared*, and *plus*, respectively) to its required operands. This is referred to as the *right-branching* [RB] *syntactic structure*. Thus the ambiguities in (1) thus result from complex NP formation. Section 2.2 describes the previous research on prosodic marking of such ambiguities.

2.2. Previous Research

For more than half a century, researchers have noted that speakers use prosodic manipulations to mark the intended meaning of syntactically ambiguous sentences and mathematical formulae. [1], working from his own intuitions, noted that the length of pauses can indicate syntactic grouping, both for mathematical and non-mathematical sentences. He gave prosodic descriptions of a wonderful example sentence from [2], reprinted here as (4), noting that by "making a longer pause next to... larger parentheses" [1:62] speakers can mark several levels of embedded structures.

- (4) Three times five minus two times two is eleven but three times five minus two times two is eighteen.

[3] and [4] likewise found that relatively longer or shorter durations of key words in complex conjoined NPs could lead to different interpretations of phrases like (5), which [5] extended to cover simple MF as in (6).

- (5) Sam and Steve or Bob will come.
 (6) a. A + (E · O)
 b. (A + E) · O

Recent work on read MF by Wagner [6,7] has used stimuli similar to (6) and shown that speakers actually produce three different types of prosodic contours in reading MF. *Left-branching prosodic structures* group early elements together via relatively smaller early prosodic breaks and a larger later break; these are used with MF like (6b). *Right-branching prosodic structures* use a larger early break and smaller later breaks to indicate structures like (6a). Speakers are also found to use *flat prosodic structures* in which all breaks are equal, but, according to [6], they only use these flat structures when they intend (6b) and are unaware of the possible ambiguity.

The use of “aware” versus “unaware” subjects in [6] is one of only two previous attempts to manipulate the level of mathematical knowledge of subjects; the other was [8]. [8] found no significant differences in productions of read MF by their “expert” and “novice” speakers, though the only criteria that made someone an “expert” reported by the authors was that these subjects were “used to listening to equations” [8:218]. Another caveat is that the data coding in this study may have obscured any differences, as only silences of greater than 300ms were noted, and even then only for 40% of their stimuli. The present study considers the mathematical skill of speakers more carefully, specifically looking for differences in prosody between more and less adept speakers.

2.3. Absolute v. Relative Boundaries

Researchers are also divided over whether the absolute strength of a particular boundary is enough to determine the desired syntactic grouping, or whether listeners must compare the relative size of two boundaries. [9] describe two general assumptions they note in the literature about they way prosodic processing works. According to the Absolute Boundary Hypothesis (ABH), the absolute strength of a boundary determines how close in syntactic structure the constituent following the boundary is to the one preceding it. According to the Relative Boundary Hypothesis (RBH), it is the strength of the boundary before a constituent relative to the strength of all preceding boundaries that determines the intended syntactic grouping. In both cases, the larger the boundary (or the larger the relative difference) the further apart constituents are in syntactic structure. [9] revisit data from [10] and perform a perception experiment, finding support for both the ABH and at least a categorical version of the RBH. A variable not considered [9] is whether relative boundary strength can vary continuously (as in the prosodic system of [11]) or only categorically (as in the Autosegmental-Metrical theory [12]).

The preceding has shown that an experiment on read mathematical formulae can contribute to understanding about the way speakers use prosody to mark the intended meaning of syntactically ambiguous strings. By looking at productions of MF containing ambiguous complex NPs by expert and non-expert subjects, this study investigates whether and how

speakers manipulate their prosody to disambiguate read MF and whether these manipulations depend on the mathematical ability of the speaker.

3. Experiment

To determine the effect of mathematical fluency on the prosody of read mathematical formulae, an experiment was conducted featuring read MF with the ambiguities discussed in § 2.1.

3.1 Subjects & Materials

Thirty undergraduate students participated in the experiment in exchange for course credit. No subjects reported vision problems. Data from four subjects was excluded – three for being non-native English speakers, one whose data was lost due to equipment failure. Thus the analysis below relies on data from 26 subjects. Subjects completed a questionnaire inquiring about their experiences with and attitudes toward mathematics. No subjects reported strong negative feelings about math (lower than three on a seven point scale). None were math, physics, or engineering majors, and none had taken any mathematics courses beyond introductory calculus.

Twenty-four pairs of target stimuli were created, eight reflecting each of the types of ambiguities shown in (3). Each stimulus was paired with a minimally different MF with an identical structure but slightly different numbers (compare 7a,b). Of these twenty-four stimuli, eight targets contained both types of ambiguities, as in (7). For stimuli that allowed multiple ambiguities, only two of the possible options were presented.

- (7) a. $\sqrt{23 - A} \cdot 2$
 b. $\sqrt{47 - A} \cdot 2$

Two experimental lists were be constructed, balancing the syntactic branching structure of each MF between subjects. 12 fillers with non-ambiguous descriptions were added to each list, so each subject saw 36 stimuli. Stimuli were presented in a pseudorandom order such that the first stimulus was a filler, minimally different stimuli that were roughly equivalent in segmental form (7a,b) did not appear with fewer than three other formulae in between, and no more than three target formulae appeared in a row without a filler. Three additional training formulae were constructed, two of which featured ambiguities in their English descriptions.

3.2 Procedure & Data Coding

Following a short questionnaire, subjects were given a set of instructions on how to read MF in the experiment, and specifically told to avoid phrases like *the quantity*, *all over*, *(open/close) parentheses*, etc., and were told not to rearrange terms of the MF. On each trial, subjects viewed one of the 36 stimulus MF on a computer monitor, generated the corresponding English utterance themselves, then spoke the utterance into a head-mounted microphone. Subjects were instructed to repeat the utterance if they caught themselves saying something incorrectly or if they were disfluent. Once finished reading, they clicked the mouse to reveal the key value of the variable A and five multiple choice answer options. They were given ten seconds to evaluate the MF

using the value of A. In some order, the five multiple choice answers always contained (1) the correct answer, (2) the incorrect answer obtained by following the other prosodic encoding (i.e., in a LB trial, the correct answer for the RB reading was always a choice) which is referred to here as the Alternate Prosody answer, and (3) three plausible incorrect answers. Subjects completed three practice trials to ensure that they understood the procedure and used the correct vocabulary to read the MF. Any productions of forbidden phrases like those above were caught by the experimenter during the practice, and the subject repeated the training phrase without the illicit wording. Subjects were reminded to speak clearly, but were not explicitly told to disambiguate. Recordings were made using Audacity and analyzed in Praat.

Utterance were transcribed using ToBI transcription criteria for break index strength. The size of the break index following the variable *A* was noted and compared to the size of earlier break indices. Since all ambiguities relied on whether the variable *A* was grouped with the preceding terms (LB reading) or the following terms (RB reading) the break following *A* can be seen as most important. Once transcribed, the pattern of break indices within utterances were compared according to several different coding schemes. Utterances were coded as *same* or *different* depending on whether the speaker's pattern of break indices was (or was not) the same on minimally different MF like (7a,b). Each trial was coded as prosodically *left-branching* if the break index after *A* was larger than any preceding index, *right-branching* if it was smaller than at least one preceding break index, and *flat* if it was equal to all other break indices following earlier operands. Utterances were coded as *matching* if they followed the predictions made in [6] (either LB or flat prosody with LB syntax, RB prosody with RB syntax) and as *non-matching* otherwise. Since it was hypothesized that flat prosody may be a separate phenomenon, non-flat prosody utterances were further coded as *cooperating* if the branching structure of the prosody matched that of the syntax and *conflicting* if it did not. Finally, flat prosodic utterances were subdivided into *big flat* and *little flat* depending on whether all relevant break indices were equal to one (little) or greater than one (big).

4. Results and Discussion

4.1 Results

Overall accuracy on the multiple-choice test was high, with an average of 84.1% (range: 58.3% - 97.2%, sd: 9.9%). Accuracy on target items was higher still, at 86.5%. Alternate Prosody incorrect answers, in which subjects chose the answer corresponding to an LB syntactic structure while trying to solve a formula with RB syntactic structure or vice versa, were exceedingly rare. These responses made on only 11 of the 83 errors on target trials, meaning that even given that subjects chose a wrong answer, they were significantly less likely to choose the Alternate Prosody answer ($p < 0.01$). MF involving the *square root* operator were the hardest, with an average accuracy of only 78.2%. compared to 87.7% for *squared*, and 93.7% for binary operators. All pair-wise t-tests were significant at $p < 0.05$ or less. Target trials involving LB syntactic structures (83.1% correct) were overall significantly harder than trials involving RB syntactic structures (89.9% correct, comparison significant at $p < 0.05$). No significant differences were observed between lists.

Of the 302 pairs of trials where subjects correctly produced utterances corresponding to both LB and RB formulae, 272 pairs (90%) were marked with *different* patterns of prosodic breaks. Subjects used a total of 198 different break index patterns across all trials. Thus even at this early stage two results are clear: speakers do use prosody to differentiate the intended meanings of segmental strings corresponding to MF, and there is no single “correct” break index pattern for a given structure.

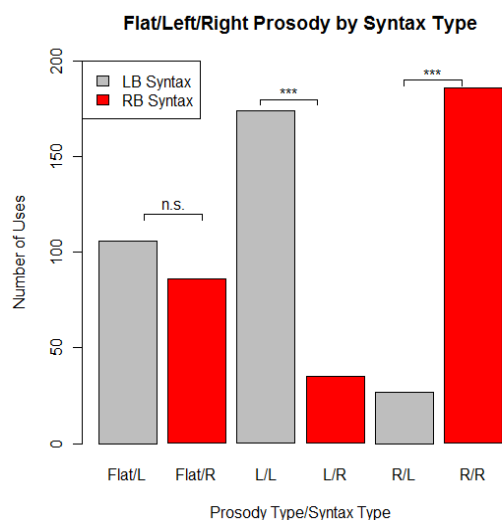


Figure 1. Use of flat/LB/RB prosody by syntax type.

Figure 1 shows the use of each type of prosody with each type of syntactic structure. Flat prosodic structures were used quite often with both types of syntactic structure, but not significantly more often with syntactically LB MF ($p > 0.02$) as predicted by [6]. Matching prosody was not used with significantly more correct answers to math problems ($p = 0.3$), nor was there a significant correlation between use of matching prosody and overall score on the multiple-choice test ($p > 0.75$). Splitting subjects at the mean score into categorical expert and non-expert groups likewise revealed no significant effect of matching versus non-matching prosody on overall test score ($p > 0.85$). Thus the *matching-non-matching* coding does not appear to be useful.

However, setting aside the prosodically flat utterances, the *cooperating-conflicting* coding of utterances is found to be useful. Cooperating prosodic structures, in which the direction of branching in prosodic and syntactic structures matched, were used in 360 of the 614 fluent trials (58.6%). Conflicting prosodic structures, in which the direction of branching mismatched, were produced in only 62 trials (10.1%), significantly fewer ($p < 0.001$). This difference between the use of cooperating and conflicting prosody shows that not only are speakers disambiguating read MF via slightly different patterns of break indices (as shown above), but that they are in fact using the left- or right-branching structure of their prosody to indicate left- and right-branching syntax. As shown in table 1, cooperating prosodic contours were more common across each problem type than flat contours, which were in turn more common than conflicting contours (all p 's < 0.01).

Table 1. Use (and accuracy) of cooperating, conflicting and flat prosody by problem type

	All	Binary Operators	Squared	Square Root
Cooperating	360 (89.4)	128 (96.0)	105 (91.4)	127 (81.1)
Conflicting	62 (75.8)	13 (92.3)	21 (66.6)	28 (75.0)
Flat	192 (84.4)	64 (89.1)	77 (88.3)	51 (72.5)

Cooperating prosody trials had significantly higher accuracy than conflicting prosody trials overall ($p < 0.05$). Use of cooperating prosody showed a significant positive correlation with item score ($p < 0.05$), while use of conflicting prosody showed a significant negative correlation with item score ($p < 0.01$). High-scoring expert subjects were also less likely to use conflicting prosody than low-scoring non-expert subjects. While the overall trend did not reach statistical significance ($p = .13$), splitting subjects at the mean score into categorical expert and non-expert groups revealed significantly more use of conflicting prosody by the non-experts than by the experts ($p < 0.05$). Thus there is some evidence that both the level of solving individual math problems and when considering an entire series of problems, subjects who read MF with cooperating prosody were more likely to solve problems correctly than those who read with conflicting prosody.

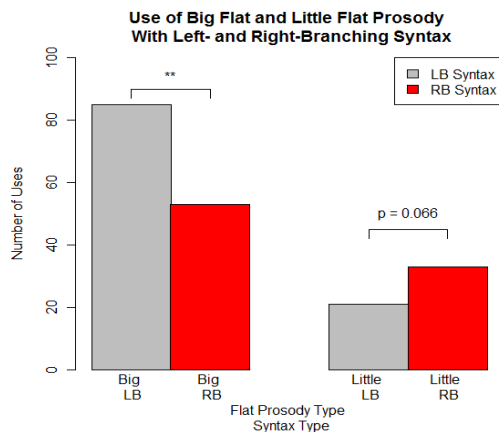


Figure 2. Breakdown of flat prosodic structures

Contrary to predictions based purely on the Relative Boundary Hypothesis, but consistent with findings from [9], there were significant differences in the meanings speakers intended when they used *little flat* and *big flat* prosody (figure 2). Little flat prosodic contours were used in 54 trials, significantly less often than big flat prosodic contours, which were used in 138 trials ($p < 0.001$). Big flat prosodic structures significantly more often indicated LB syntax than RB syntax ($p < 0.01$), while little flat prosodic structures more often indicated RB syntax than LB syntax, though this final comparison just misses significance ($p = 0.066$).

4.2 Discussion

These results support a prosodic system including both the Relative Boundary Hypothesis and the Absolute Boundary Hypothesis. The relative size of two break indices in the complex NPs studied here mattered when the indices were different: operands set off by larger break indices were intended to apply later in the process of constructing the complex NP. In trials with RB syntactic structure, subjects occasionally used prosodic structures where an intermediate phrase break appeared after the variable *A* and after an earlier intonational phrase break, creating RB prosodic structures that matched the syntax. If we were forced to rely on the ABH alone, these structures would be incorrectly coded as LB, since a substantial break appeared after the variable. This coding would miss the fact that the earlier major break allowed the reader to convey the intended structure. When two break indices were equal, the absolute size of those indices mattered: equally sized large breaks in the big flat condition indicated RB syntactic structures, but equally sized small breaks in the little flat condition were used to indicate LB syntactic structures. Thus, both relative boundary size and absolute boundary size are important. If we had to rely only on the RBH, as would be required by [11]'s model, we would not be able to distinguish these types of utterances.

Finally, significant differences were found between the use of cooperating and conflicting prosody by more and less mathematically adept speakers. This is contrary to earlier findings [8], but not necessarily surprising given that more adept mathematicians may have more experience with MF and small differences in MF may be more salient to these subjects.

5. References

- [1] Pike, K. *The Intonation of American English*. Ann Arbor, MI: University of Michigan Press 1945.
- [2] Bloomfield, L. *Linguistic Aspects of Science*. Chicago, IL: University of Chicago Press, 1939.
- [3] Lehiste, I. "Phonetic disambiguation of syntactic ambiguity." *Glossa* Vol. 7, 107-122. 1973.
- [4] Lehiste, I., Olive, J., & Streeter, L. "Role of duration in disambiguating syntactically ambiguous sentences." *Journal of the Acoustical Society of America*, Vol. 60, 1199-1202. 1976.
- [5] Streeter, L. "Acoustic determinants of phrase boundary perception." *Journal of the Acoustical Society of America*, Vol. 64, 1582-1592. 1978.
- [6] Wagner, M. "Prosodic options or syntactic/semantic choices?" presented at the Workshop on Prosody and Meaning. Barcelona, Spain, 2009.
- [7] Wagner, M., & Crivellaro, S. "Relative Prosodic Boundary Strength and Prior Bias in Disambiguation." *Speech Prosody*, University of Chicago. 2010.
- [8] O'Malley, M., Kloker, D., & Dara-Abrams, B. "Recovering parentheses from spoken algebraic expressions." *IEEE Transactions on Audio and Electroacoustics*, Vol. 21 No. 3, 217-220, 1973.
- [9] Snedeker, J., & Casserly, E. "s it all relative? Effects of prosodic boundaries on the comprehension and production of attachment ambiguities." *Language and Cognitive Processes*, Vol. 25, 1234-1264, 2010.
- [10] Snedeker, J. & Trueswell, J. "Using prosody to avoid ambiguity: Effects of speaker awareness and referential context." *Journal of Memory and Language*, Vol. 48, 103-130, 2003.
- [11] Wagner, M. *Prosody and Recursion*. Ph.D. Dissertation, Massachusetts Institute of Technology, MA, 2005.
- [12] Beckman, M. & Pierrehumbert, J. "Intonational structure in Japanese and English." *Phonology Yearbook*, Vol 3, 255-309, 1986.