

PROSODIC STRUCTURE, PERFORMANCE STRUCTURE AND PHRASE STRUCTURE

Steven Abney

Bell Communications Research
445 South St., Morristown, NJ 07962-1910

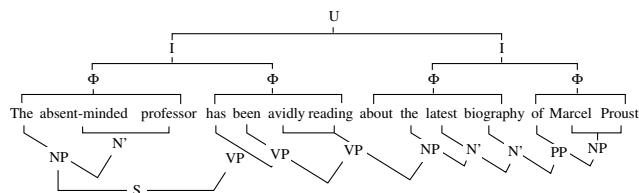
ABSTRACT

It is natural to expect phrase structure to be important in predicting prosodic phrasing. Yet there appears to be a consensus that syntactic phrases do not correspond well to prosodic phrasing, and independent structures have been proposed to account for prosody.

I propose that the problem with phrase structure lies with the particular measures of boundary strength applied to syntactic structures, and with the fact that phrase structure is viewed as an immediate constituency tree exclusively. I propose viewing phrase structure as a composite of immediate constituency and dependency relations, and present an alternative measure of boundary strength. I show that boundary strength according to this measure corresponds much more closely to empirical prosodic (and psycholinguistic) boundary strength than does syntactic boundary strength according to a standard measure.

1. INTRODUCTION

It is natural to expect phrase structure to be important in predicting prosodic phrasing. Hence it is somewhat unsettling that the relationship between prosodic and syntactic structure appears so tenuous, as for example in Selkirk's account [10, 11, 12]. Selkirk's prosodic structure differs from standard phrase structure on several counts, but most notably because it is much flatter than standard phrase structure, which is heavily right-branching in English:



In a similar vein, some psycholinguists have concluded that syntactic structure provides an inadequate model of

the *performance structures* reflected in linguistic behavior. Martin, Grosjean, and others have explored experimental measures of the relative prominences of boundaries between words, and conclude that the syntactic prominence of a boundary is not the best predictor of its empirical prominence [4, 6, 7, 8, 9].

If prosodic structures and performance structures differ from phrase structure, however, they appear to correspond well to each other. For example, Gee and Grosjean [6] use Selkirk's prosodic phrases in an algorithmic model of their experimental data. And turnabout being fair play, Bachenko and Fitzpatrick [3] adapt Gee and Grosjean's algorithm to predict prosodic structure for speech synthesis.

However, I believe the perceived inadequacy of syntactic structure is at least in part an artifact of measures of syntactic boundary prominence that are based on immediate-constituency trees alone. I would like to show that we can obtain a measure of syntactic boundary prominence that corresponds better to prosodic and psycholinguistic boundary prominence if we view phrase structure as a composite of immediate constituency and dependency relations.

2. CHUNKS AND DEPENDENCIES

I propose that the structure relevant for prosody and performance is a composite of immediate-constituency and dependency relations. Usually, dependency grammar is an alternative for representing phrase structure, in competition with immediate constituency. However, there is often a systematic correspondence between dependencies and immediate constituency. I will assume such a correspondence, and *define* dependency in terms of immediate constituency, as follows:

Y depends on X iff
X is a word, and
Y is an immediate constituent of a phrase headed by X

Graphically:



Dependencies are combined with immediate constituency in the relation *is licensed by*. X may license Y either by dependency or by immediate constituency:

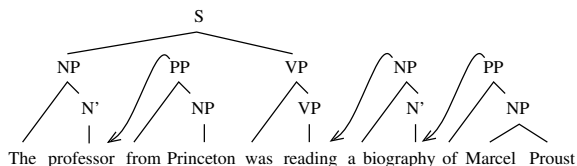
X licenses Y by dependency iff
 Y depends on X, and
 X is a major-category head
 (N, V, Adj, or Adv), and
 X precedes Y

X licenses Y by immediate constituency iff
 Y is an immediate constituent of X, and
 there is no node that licenses Y by dependency

Consider, for example, the following sentence (adapted from [10]):

the absent-minded professor from Princeton
 was reading a biography of Marcel Proust

The major-category heads are *absent-minded*, *professor*, *Princeton*, *reading*, *biography*, *Marcel Proust*. The PP *from Princeton* follows, and depends on, *professor*; hence *from Princeton* is licensed by dependency. Likewise for *a biography* (depends on *reading*), and *of Marcel Proust* (depends on *biography*). These three phrases are licensed by dependency; all the other phrases are licensed by immediate constituency. We can represent the licensing structure as follows, where the arrows represent licensing by dependency, and the straight lines represent licensing by immediate constituency:

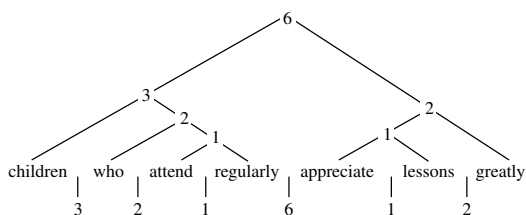


There is a certain similarity between this structure and Selkirk’s prosodic structure. In particular, if we consider only the relation *licenses by immediate constituency*, and excise the clausal node (S), the remaining connected pieces of phrase structure—which I call *chunks*—are Selkirk’s ϕ -phrases. Gee and Grosjean also base their algorithm on ϕ -phrases. The correspondence between chunks and ϕ -phrases suggests that licensing structure

might do better than standard phrase structure in predicting prosodic and performance-structure boundary prominence.¹

3. MEASURING SYNTACTIC BOUNDARY STRENGTH

Given phrase structure trees, we also require a method for computing boundary prominence. The method that I take to be “standard” is the one assumed in the performance-structure literature, by which the prominence of a boundary *b* is the number of non-terminal nodes in the smallest constituent spanning *b*. For example:



I would like to propose an alternative measure. The general idea is as follows:

1. Clause boundaries > chunk boundaries > word boundaries
2. “Strong” dependencies between immediately adjacent chunks/clauses weakens the boundary between them
3. Phonologically weak chunks “cliticize” to the adjacent chunk

Phonologically weak chunks are chunks containing a single word whose category is pronoun, particle, auxiliary, conjunction, or complementizer. The following are specific boundaries weakened by “cliticization”:

- verb – (indirect) object pronoun
- verb – particle
- subject pronoun – verb
- wh* pronoun – auxiliary
- inverted auxiliary – subject
- conjunction – subject
- complementizer – subject

¹An analysis in which phrase structure consists of a series of strata—words, chunks, simplex clauses—also proves useful for rapid, robust parsing of unrestricted text [1, 5]. The parsing advantages of chunks provided my original motivation for considering them. I undertook the work described here in order to make good on earlier hand-waving about a possible relation between chunks and prosodic phrases.

The “strong” dependencies are these:

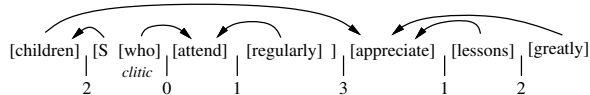
- verb – any dependent
- noun – *of* phrase
- noun – restrictive relative clause
- subject – verb

I also relax the adjacency requirement to permit one intervening phonologically weak chunk. In particular, if a particle or indirect object pronoun intervenes between a verb and its following dependent, the boundary before the dependent is still weakened.

I assign the following heuristic values to boundaries. What is important for my purposes is the relative values, not the absolute values.

- 3 Unweakened clause boundary
- 2 Unweakened chunk boundary
- 2 Weakened clause boundary, governor is noun
- 1 Weakened clause boundary, governor is verb
- 1 Weakened chunk boundary
- 0 Weakened chunk boundary involving phonologically weak chunk
- 0 Intra-chunk word boundaries

To illustrate the measure, consider the following example from Martin [9]:



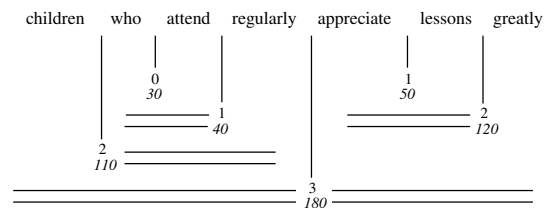
The bold arrows mark dependencies that induce weakening. The first boundary is a clause boundary, weakened from 3 to 2. The second boundary is a chunk boundary. Since *who* is phonologically weak, the boundary is weakened to 0. The third boundary is a chunk boundary weakened from 2 to 1. The fourth boundary is an unweakened clause boundary, value 3. The next to last boundary is a weakened chunk boundary, and the final boundary is an unweakened chunk boundary.

4. COMPARING THE MODELS

To compare the chunks-and-dependencies model to the standard model, we need to compare both models to boundary-prominence data. I am primarily interested in the local relative prominence of boundaries. A boundary *b* is defined to be *locally more prominent* than boundary *c* iff *b* is more prominent than *c* and every intervening

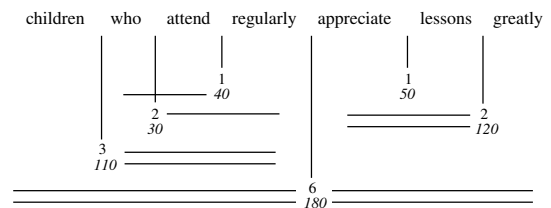
boundary is less prominent than *c*. In comparing theoretical and empirical prominences, each *inversion* counts as an error. An inversion arises wherever *b* is locally more prominent than *c* according to the model, but *c* is locally more prominent than *b* according to the data.

To illustrate, consider again sentence (1), with theoretical and empirical boundary prominences:



The top numbers are the boundary prominences according to the chunks-and-dependencies model; the bottom numbers (in italic) are empirical values obtained by Martin [9] in a naive-parsing experiment. The length of the vertical lines corresponds to the theoretical prominence of each boundary. The horizontal lines represent the local relative prominence domain of each boundary: the solid lines according to the model, the dotted lines according to the data. In this case, the theoretical and empirical domains match exactly.

This is the same sentence, using the standard model:



In this case, there is an inversion: the second boundary is more prominent than the third, according to the model, but the third is more prominent than the second, according to the data. The inversion is reflected in the line crossing.

5. DATA

To compare the models, I examined two sets of data: performance structure data reported by Grosjean, Grosjean and Lane [7]; and a set of sentences with hand-marked prosodic boundaries, kindly provided by Julia Hirschberg of AT&T Bell Laboratories.

Grosjean, Grosjean and Lane conducted two experiments, one examining pauses when subjects read sentences at various speeds, and one examining parsing by

linguistically-naive subjects. They report only the data on the pausing experiment, though they claim that the parsing data is highly correlated with the pausing data.

The data consists of 14 sentences, containing 55 opportunities for inversions. (An opportunity for inversion is a boundary that, according to the model, is locally more prominent than at least one other boundary). In 52 cases the model makes the correct prediction (5% error). The three inversions all involved unexpectedly prominent boundaries around multisyllabic pre-head modifiers at sentence end, hence they arguably reflect a single unmodelled effect. Using the standard measure gives us 42 inversions out of 102 opportunities for inversion, or 41% error, dramatically worse than the licensing measure's 5% error rate. (There are more opportunities for inversion because the standard model typically makes more distinctions in boundary prominence.)

The second data set consists of 127 sentences from the Darpa ATIS task, with prosodic boundary markings added by Julia Hirschberg. She distinguished three boundary strengths: strong, weak, and no boundary.

A complication in the prosodic data is the presence of hesitation pauses, which I do not expect a syntactic model to capture. As a primitive expedient, I formulated a rule that I could apply mechanically to distinguish hesitation pauses from "genuine" prosodic boundaries, and I eliminated those boundaries that were hesitation pauses according to the rule. Namely, I eliminated any prosodic boundary immediately following a preposition, conjunction, infinitival *to*, or a prenominal modifier.

After eliminating hesitation pauses, I applied the licensing-structure measure and the standard measure. Using the licensing measure, there were 363 opportunities for inversions, and 12 observed (3% error). Applying the standard model to 16 sentences drawn at random from the data gives 38 inversions out of 114 opportunities, or 33% error.

Caution is in order in interpreting these results, in that I have not controlled for all factors that may be relevant. For example, the standard measure generally has a greater range of distinctions in boundary prominence, and that may lead to a larger proportion of errors. Also, the method I use to eliminate hesitation boundaries may help the chunks-and-dependencies model more than it helps the standard model. In short, these are exploratory, rather than definitive results. Nonetheless, they strongly suggest that the chunks-and-dependencies model corresponds to empirical prominences better than the standard model does, hence that syntactic structure may be a better predictor of prosodic and performance

structures than previously thought.

References

1. Steven Abney. Rapid Incremental Parsing with Repair. *Proceedings of the 6th New OED Conference: Electronic Text Research*. University of Waterloo, Waterloo, Ontario. 1990.
2. Steven Abney. Chunks and Dependencies: Bringing Processing Evidence to Bear on Syntax. Paper presented at the Workshop on Linguistics and Computation, University of Illinois, Urbana/Champaign. 1991.
3. J. Bachenko & E. Fitzpatrick. A Computational Grammar of Discourse-Neutral Prosodic Phrasing in English. *Computational Linguistics* **16**(3), 155-170. 1990.
4. Jean-Yves Dommergues & François Grosjean. Performance structures in the recall of sentences. *Memory & Cognition* **9**(5), 478-486. 1981.
5. Eva Ejerhed. Finitary and Stochastic Methods of Clause Parsing.
6. James Paul Gee & François Grosjean. Performance Structures: A Psycholinguistic and Linguistic Appraisal. *Cognitive Psychology* **15**, 411-458. 1983.
7. F. Grosjean, L. Grosjean, & H. Lane. The patterns of silence: Performance structures in sentence production. *Cognitive Psychology* **11**, 58-81. 1979.
8. W.J.M. Levelt. Hierarchical chunking in sentence processing. *Perception & Psychophysics* **8**(2), 99-103. 1970.
9. Edwin Martin. Toward an analysis of subjective phrase structure. *Psychological Bulletin* **74**(3), 153-166. 1970.
10. Elisabeth O. Selkirk. On prosodic structure and its relation to syntactic structure. In T. Fretheim (ed.), *Nordic Prosody II*. Tapir, Trondheim. 1978.
11. Elisabeth O. Selkirk. Prosodic Domains in Phonology: Sanskrit Revisited. In M. Aronoff and M.-L. Kean (eds.), *Juncture*. Anma Libri, Saratoga, CA. 1980.
12. Elisabeth O. Selkirk. On the Nature of Phonological Representations. In T. Myers, J. Laver, J. Anderson (eds.), *The Cognitive Representation of Speech*. North-Holland Publishing Company, Amsterdam. 1981.
13. Elisabeth O. Selkirk. *Phonology and Syntax: The Relation between Sound and Structure*. The MIT Press, Cambridge, MA. 1984.