PERCEPTION OF PROMINENCE AND BOUNDARIES BY NAÏVE FRENCH LISTENERS

Caroline L. Smith

University of New Mexico, USA
caroline@unm.edu

ABSTRACT

French phrasal structure is described as based on an accent group in which the final (non-schwa) vowel receives prominence by virtue of its position in the phrase. This co-occurrence of prominence and phrasal boundaries was verified experimentally by testing the perceptions of untrained participants. The listeners demonstrated a strong tendency to mark as prominent those words immediately preceding locations perceived as boundaries. Analyses of acoustic and syntactic properties of these locations are consistent with the idea that listeners used similar strategies in labeling both prominent words and boundaries.

Keywords: French, perception of prosody, spontaneous speech

1. INTRODUCTION

Although the prosody of spoken French has been studied extensively (see, e.g., [5, 6] for reviews), relatively little work has investigated ordinary listeners' understanding of prosodic organization (for exceptions, [7, 8, 9]). The study reported here applies a method previously applied to other languages, notably English ([2, 3]), as a means to investigate naïve listeners' perceptions of spontaneous spoken French.

Here the focus is the smallest phrasal unit in French, the accent group, described as having prominence on the last full-vowel syllable [4]. That is, the prominence immediately precedes the phrasal boundary. Testing whether this description accords with listener judgments, and examining these in conjunction with the acoustic and syntactic properties of the speech, can inform us about the factors that guide the perception of prosody.

2. METHOD

2.1. Materials for the listening experiment

Two types of speech materials were used. One set of ten extracts was prepared from recordings of a map task experiment that had been previously recorded at a Paris university [11]. The speakers are ten female undergraduates from the Paris region. They were recorded individually in a task which required them to give directions over the telephone as to how to use the Paris métro system to travel to various destinations. These extracts consist of fairly informal, spontaneous task-directed speech. The extracts were selected from portions of the conversations during which the one speaker had a relatively long conversational turn, without overlap by the interlocutor. These extracts varied from 13 to 24 seconds in length.

The second set of ten extracts was taken from a discussion of television advertising that was broadcast on a current affairs program on the France Info radio station. These extracts also consist of single-speaker passages of spontaneous conversational speech, but the speakers are journalists and public figures. Their conversation was recorded for broadcast and thus illustrates a more formal register. The selected samples include two extracts from each of the five speakers who participated in the discussion. The duration of these extracts is from 26 to 53 seconds.

Orthographic transcriptions of the extracts were prepared by the experimenter (a fluent non-native speaker of French), then edited by a native speaker with phonetic training. These transcriptions were prepared for use in the listening test by removing punctuation and line breaks except as necessary to fit on the page, in order to avoid providing any hints as to the structure. Disfluencies such as repeated or partial words were included in the transcripts but filled pauses ("euh…") were not indicated. Three additional extracts were also prepared to serve as practice samples.

2.2. Participants and testing procedure

Fifty-one listeners without advanced training in phonetics or prosody were recruited at three higher education institutions in France. Most were undergraduate students in linguistics. In order to test listeners in groups for efficiency reasons, they were not screened for native language, and thus, a

few were included who are non-native speakers of French. Different listeners participated in the experiment in different settings: some were tested in groups of 5-17 in a classroom, others individually or in groups in a sound-attenuated room. Each listener was presented with a packet containing instructions and the printed transcripts of the practice and test extracts. They marked their responses on these print-outs.

Listeners were randomly assigned to one of two groups. 26 listeners were asked to mark a vertical line between words at locations where they perceived a boundary between two phrasal units (*syntagmes*, defined as groups of words that form a single unit for both meaning and function).[1] The remaining 25 listeners were instructed to underline words that were highlighted (*mis en relief*). This instruction was given even though prominence in French is associated with syllables, as the need to respond rapidly would make it too hard to mark single syllables. All listeners heard the extracts in the same order, with brief pauses between them controlled by the experimenter depending on the listener(s)' wishes. They practiced first on two practice map task extracts, then responded to ten map task extracts, then practiced on an extra broadcast extract, then responded to the ten test broadcast extracts. The extracts in each group were presented in random order. No two extracts with the same speaker were presented consecutively.

One listener in the boundary-marking group failed to follow directions, so that individual's responses were excluded from analysis, leaving a total of 25 listeners in each group. All of their responses to the twenty test extracts were retained, and coded in Excel spreadsheets.

2.3.   Statistical analyses

Agreement among listeners was assessed using a modified form of Cohen's Kappa, which takes into account the amount of agreement that can be expected by chance. Kappa values can vary between 0 and 1. The particular form of Kappa used here is based on [1]. Calculations were made using the Online Kappa Calculator [10]. Kappa values were determined for each extract, pooling across all the listeners.

Most of the results reported here are counts and correlations that were calculated in Excel. These include calculation of a boundary score for each word, equal to the proportion of listeners who marked that word as followed by a boundary.

Those words marked by two-thirds or more of listeners (17 or more of the 25) were considered to have "consensus" agreement. This criterion was arbitrary but indicates a substantial consensus.

2.4.   Part of speech and sentence-end labeling

All the words were automatically labeled for part of speech using [13]. These labels were then reviewed and hand-corrected as necessary. The experimenter also manually labeled each point in the extracts that was judged to be a possible end of a sentence. Because the speech being analyzed was produced spontaneously, most of it does not consist of grammatically complete sentences. The principal criterion for labeling a location as a possible sentence end was if the utterance would be readily interpretable if it ended at that location. In most cases this meant that all obligatory complements for the verb were present by that point in the utterance.

3.   RESULTS

The goal of this study was to examine the extent to which perceived prominences and boundaries co-occurred, and to see whether listeners were using similar information to identify them. Before using the listeners' responses as the basis for analyses, the rates of agreement among listeners were calculated to ensure that they are sufficiently high.

3.1.   Rates of agreement among listeners

The values of the kappa statistic used to assess agreement ranged from .53 to .80 for marking of prominence, with a mean of .69 across the 20 extracts. For boundary marking, kappa ranged from .75 to .88 with a mean of .83. Randolph [10] suggests that for this form of kappa, .7 or above is "adequate", which means that the agreement for prominence marking is borderline, but the rate for boundaries is well above this proposed cut-off. Boundaries identified on the basis of the listeners' responses can thus be taken with confidence, while the locations of prominences are less reliable.

3.2.   Distribution of prominent words and boundaries perceived by listeners

As described in section 2.3, "consensus" markings were identified as those locations where at least 2/3 of the listeners had marked a boundary or prominence. The mean number of marked locations per extract is given in Table 1.

Table 1: Mean count per extract of prominent words and boundaries identified by at least 2/3 of the listeners.

	Map tasks	Broadcast
Total words per extract	55.9	134.8
Prominent words	2.2	4.3
Boundaries	3.5	8.5

Listeners marked boundaries less often than prominences. The global median was one boundary marked every 9.7 words, and one prominence every 8.6 words. Prominence scores of words before and after consensus boundaries were examined in order to determine whether the words in these positions have prominence scores that diverge from the average. Table 2 shows that, as expected, words before boundaries received much higher prominence scores than the average. Words after boundaries received lower prominence scores than the average.

Table 2: Mean prominence scores for all words, and words before and after consensus boundaries.

	Map tasks	Broadcast
All words	0.16	0.11
Words before boundaries	0.65	0.42
Words after boundaries	0.08	0.09

In order to further investigate the relation between the listeners' marking of boundaries and of prominent words, the correlation was calculated between the prominence scores and boundary scores of the words in each extract. This analysis did not examine specific boundary locations; rather, it looked at the overall relation between locations marked as boundaries and the prominence of words before them. Because prominences were marked more frequently than boundaries, there cannot be a perfect correlation between them.

Word prominence and a following boundary had an average correlation of .68 with a standard deviation of .08 across the ten map task extracts. All ten of these are significant at $p<.001$. Across the ten broadcast extracts, correlations averaged .51 with a standard deviation of .17. These are also significant at $p<.001$, with one exception where $p=.001$. These results support the hypothesis that listeners tend to perceive prominence on words where they perceive a boundary following.

3.3. Acoustic cues to prosodic structure

If listeners are perceiving prominences and boundaries at the same locations, then these should be cued by similar acoustic features. Analyses

reported elsewhere [12] have shown that pauses (of 150 ms or longer) and F0 movement were strong cues to boundary perception in these extracts. Pauses also cued prominence. In the map task extracts, 16 of 18 words perceived as prominent preceded a pause; in the broadcast extracts, 15 of 25 were pre-pausal.

3.4. Parts of speech

As a further test to see whether similar types of words were marked by listeners as prominent and as preceding a boundary, part of speech was examined. Given that accentual groups in French are described as ending with a lexical word, it was expected that only lexical words could be perceived as prominent or before a boundary.

Figure 1: Map task extracts: Percent of words of each type with consensus labeling as prominent or as followed by a boundary.
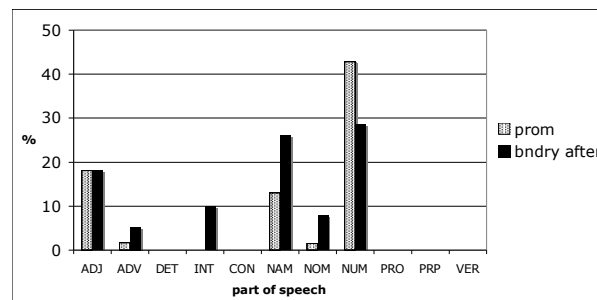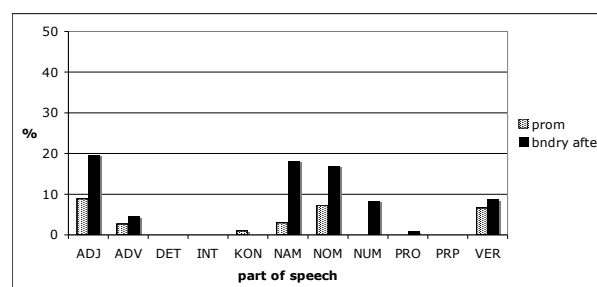

Figure 2: Broadcast extracts: Percent of words of each type with consensus labeling as prominent or as followed by a boundary.

Figures 1 and 2 show that the words most often marked by listeners as prominent or pre-boundary were nouns (NOM), adjectives (ADJ), proper names (NAM) or numbers (NUM). The high proportion of numerals and names in the map task extracts is explained by their use in references to metro lines by number and metro stations by name. No verbs (VER) were marked in the map task extracts, and a low percentage of verbs were marked in the broadcast extracts. This probably reflects the high proportion of auxiliaries among the verbs; since auxiliaries would not be expected to end an accent

group, they are unlikely to be perceived as prominent or pre-boundary.

Parts of speech that were more often perceived as prominent were also more likely to be perceived as pre-boundary, even though fewer words were marked as prominent than as pre-boundary.

The most notable exception was the higher proportion of numbers perceived as prominent than as pre-boundary in the map tasks. These occurred in contexts such as shown below, where *deux* 'two' was perceived as prominent but not pre-boundary.

vous prenez le métro deux direction Porte Dauphine.
'you take the metro two towards Porte Dauphine'

3.5. Possible sentence boundaries

Listeners' responses were expected to reflect syntax as well as acoustic prosodic cues. Locations that were judged by the experimenter to be a potential end of a sentence grammatically were much more likely to be perceived as a boundary by the listeners. In the map tasks, the word preceding a possible sentence end was also more likely to be perceived as prominent, but this was not true for the broadcast extracts.

Table 3: Total number of words in each set of extracts for which perceived prominent words and boundaries did or did not coincide with a possible sentence end.

	Map tasks	Broadcast
Precedes boundary, possible sentence end	35	69
Possible sentence end, no boundary	0	13
Prominent word, possible sentence end	21	17
Possible sentence end, not prominent	1	23

Sentence structure appears to have had a more reliable effect on boundary perception than perception of prominence. Despite the high correlations reported in section 3.2, listeners may access some different information in deciding on the locations of boundaries and prominences.

4. DISCUSSION

The precise status of the boundaries that listeners perceived is an open question, as it is impossible to know how they interpreted the instructions. Syntactic structure was clearly influential; however, the sentence fragments and disfluencies typical of spontaneous speech, found in abundance in these materials, suggest that identification of syntactic structure would have been difficult. It is likely that listeners responded to some

combination of readily-noticeable syntactic breaks and overt acoustic cues to boundaries. Although sentence structure seems to have played less of a role in perception of prominence, the traditional description for French as having a tight linkage between the occurrence of boundaries and prominences is well-justified by the findings here.

5. REFERENCES

[1] Brennan, R., Prediger, D. 1981. Coefficient Kappa: Some uses, misuses and alternatives. *Educational and Psychological Measurement* 41, 687-699.

[2] Cole, J., Mo, Y., Baek, S. 2010. The role of syntactic structure in guiding prosody perception with ordinary listeners and everyday speech. *Proc. Lang. and Cog.,* 25, 1141-1177.

[3] Cole, J., Mo, Y., Hasegawa-Johnson, M. 2010. Signal-based and expectation-based factors in the perception of prosodic prominence. *Laboratory Phonology* 1, 425-452.

[4] di Cristo, A. 2000. Vers une modélisation de l'accentuation en français (seconde partie). *J. French Language Studies* 10, 27-44.

[5] di Cristo, A. 2005. Éléments de prosodie. In: Nguyen, N., Wauquier-Gravelines, S., Durand, J. (eds), *Phonologie et Phonétique : forme et Substance*. Paris: Lavoisier, 117-157.

[6] Lacheret-Dujour, A., Beaugendre, F. 1999. *La prosodie du Français*. Paris: CNRS Editions.

[7] Obin, N., Rodet, X., Lacheret-Dujour, A. 2008. French prominence: a probabilistic framework. *Proc. ICASSP 2008* Las Vegas, 3993-3996.

[8] Pagel, V., Carbonell, N., Laprie, Y., Vaissière, J. 1995. Spotting prosodic boundaries in continuous speech in French. *Proc. 13th ICPhS* Stockholm, 4, 308-311.

[9] Portes, C. 2000. *Approche du Rôle de la Prosodie Dans la Structuration du Discours Oral en Français*. DEA thesis, Université de Provence.

[10] Randolph, J. 2008. Online Kappa Calculator. *http://justus.randolph.name/kappa*

[11] Smith, C. 2007. Prosodic accommodation by French speakers to a non-native interlocutor. *Proc. 16th ICPhS* Saarbrücken, 1081-1084.

[12] Smith, C. To appear. Acoustic correlates of listener-identified boundaries in spontaneous French speech. *Proc. of LARP 2010*, Cascadilla Proceedings Project.

[13] Tree Tagger. *http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/DecisionTreeTagger.html*

---

[1] The drawback to using the term *syntagme* in the instructions is that it can imply a syntactic unit. However, it also means "phrase", both syntactically and prosodically, and there is no other term that seems a better alternative.