Studying French mapping of syntax to prosody in natural speech

Fabián Santiago, Camille Dutrey, Martine Adda-Decker
LPP (UMR 7018) - Université Sorbonne Nouvelle & CNRS, Paris, France
LIMSI - CNRS, Université Paris-Saclay, France
{fabian.santiago-vargas, madda}@univ-paris3.fr, {madda, camille.dutrey}@limsi.fr

Background. The minor prosodic constituent in French is the so-called 'Groupe Accentuel' or Accentual Phrase (AP). This unit is often defined by a mapping between the morpho-syntactic and prosodic structures. According to this approach, in French, an AP consists in minimally a lexical word and all dependent words positioned at its left side [3, 5, 7]. The rightmost full syllable of the last lexical word must carry a pitch movement (mainly a rising pitch accent) marking the AP's right boundary. However, it is well known that beside syntactic structure, other factors contribute to the realization of APs: the number of syllables (3/4 on average) [6], articulation rate [8] and rhythm (balance of AP's size in terms of number of syllables) [2].

Research Goals. Studies analyzing the relation between syntactic and prosodic structure by using large corpora and semi-automatic acoustical analysis are still underrepresented. In this contribution, we propose to analyse the syntactic-prosodic mapping for analysing the formation of APs in French by using rich and large corpora of natural speech. Our goal is twofold: (i) studying whether grammatical information is a good predictor for the formation of APs in French in natural speech and (ii) analysing the contribution of constituent length (number of syllables within a phrase), constituent type and temporal cues to the realization of these prosodic units.

Methods. We use a subset of the French ETAPE corpus compiled by [4], which has been grammatically annotated in the French Spoken Treebank (FST) by [1]. The corpus consists of four radio broadcasts emphasizing spontaneous conversational speech (25 speakers, 1.5 hours and 16,377 words). Syntactic labels contained in the FST (fig. 1) were semi-automatically extracted and associated to the signal as Praat's TextsGrids (fig. 2). In the current study, we focus on the most frequent syntactic constituents observed in the corpus: noun phrases, verbal nuclei and adjectival phrases. Syntactic structures have few internal structure : determiners and modifying adjectives at the same level in the noun/adjectival phrases, and clitics, auxiliaries, negation and verb in the verbal nuclei. APs were predicted using morpho-syntactic rules: any lexical word and its dependent function words at its left side call for a final prosodic boundary. Predicted AP boundaries were checked against the measured prosodic boundary cues as follows: an AP was considered intonationally marked when its last metrically strong syllable was associated with a relevant rising pitch accent > 2 st. (fig 2) (falling patterns were discarded from the analysis). Pitch values (differences in st. between the final and initial vowel of the last constituent word) were measured in order to detect pitch accents marking final AP's boundaries. In addition, normalized word and constituent durations (duration of word or constituent / number of comprising phones in the word or constituent) were computed.

Results and conclusion. The syntactic-prosodic mapping described above led us to identify 4,415 potential APs in our data classified as follows: 2,528 noun phrases, 1,713 verbal nuclei and 428 adjectival phrases. However, the acoustic measurements led us to identify 1,901 APs: only 43% of the expected APs were actually detected in the signal by the presence of a pitch accent. In order to find explanations for the poor effectiveness of the mapping rules, we considered those factors that may affect the distribution of APs in our data. The first concerns the acoustic cues measured: the pitch accent criterion (> 2 st) may be too selective and may not detect AP's boundaries marked by temporal cues such as final lengthening. The second one concerns the rhythmic effects: more than 40% of our data consists of monosyllabic constituents, which do not tend to be produced as independent APs in the signal. The third one concerns the constituent type: results show that syntactic-prosodic mapping is less accurate for verbal nuclei compared to the rest of categories (fig. 3), i.e. pitch accents are less frequent in verbal nuclei than in the rest of categories. In addition, verbal nuclei tend to have shorter durations and are produced with smaller final pitch excursions (fig. 4). We thus suggest that apart from rhythmic effects (constituent size), prosodic features inherent to verbal nuclei could play a role in the formation of APs in spontaneous speech.

References

1. Abeillé A. & Crabbé B. (2013), Vers un treebank du français parlé. In *Proceedings of TALN*, 2013, pp. 174–187.
2. Delais-Roussarie, E. (2000), Vers une nouvelle approche de la structure prosodique, *Langue Française 126,* 92-112.
3. Di Cristo A. (1998), Intonation in French, in Hirst D. & Di Cristo A. (eds), *Intonation Systems: A survey of twenty Languages*. Cambridge: Cambridge University Press, pp. 195–218.
4. Gravier G., Adda G., Paulsson, N. Carré, M., Giraudel A. & Galibert O. (2012), The ETAPE corpus for the evaluation of speech-based TV content processing in the french language, in *Proceedings of LREC,* 2012, pp. 114–118.
5. Jun, S.-A. & Fougeron C. (2002), The Realizations of the Accentual Phrase in French Intonation, *Probus 14*, pp. 147–172.
6. Martin, P. (1987). Prosodic and rhythmic structure in French. *Linguistics 5*(5), pp. 925-949.
7. Post B. (2000), *Tonal and phrasal structures in French intonation*. The Hague: Holland Academic Graphics.
8. Post B. (2011), The multi-faceted relation between phrasing and intonation in French. In Gabriel C. & Ll. Conxita (eds), *Intonational Phrasing at Interfaces: Cross-Linguistic and Bilingual Studies in Romance and Germanic*, Amsterdam: John Benjamins, 2011, pp. 43–74.
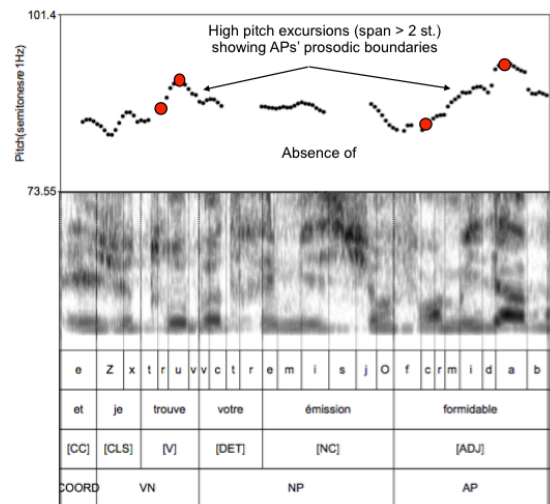
Figures
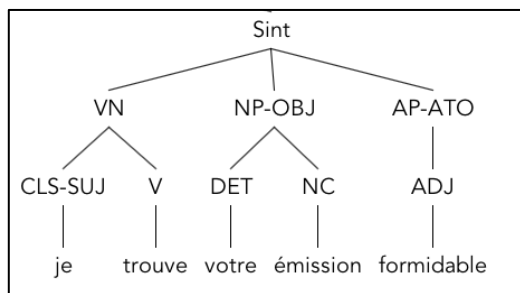
Fig. 1. Grammatical annotations of the clause *je trouve votre émission formidable* ('I found (that) your podcast (is) great') provided by the French Spoken Treebank.

Fig. 2. Spectrograme, f0 trace and grammatical annotations of the clause *je trouve votre émission formidable* ('I found (that) your podcast (is) great').
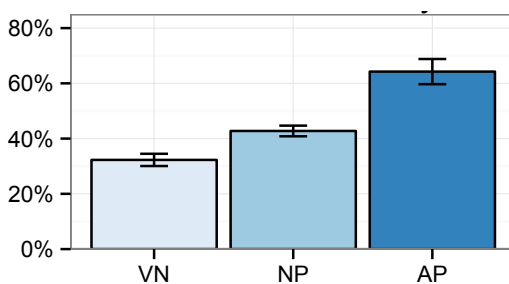
Fig. 3. Proportions of boundary tones across the three constituent types : verbal nuclei (VN), noun phrases (NP) and adjectival phrases (AP).
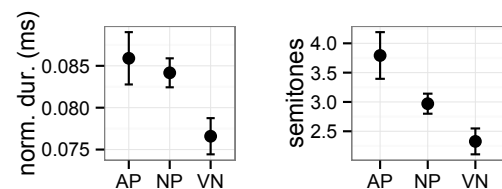
Fig. 4. Durations (left) and final f0 delta values (right) across the three constituent types : verbal nuclei (VN), noun phrases (NP) and adjectival phrases (AP).