

## Editorial

(Selected papers from the 2013 Conference on the Prosody-Discourse Interface)

**Piet Mertens\* & Anne Catherine Simon\*\*<sup>1</sup>**

\* Linguistics Department, University of Leuven, Belgium

\*\* Institut Langage et Communication, Université Catholique de Louvain, Louvain-la-Neuve, Belgique

---

This issue of JoSS contains extended versions of a small selection of papers presented at the *Prosody-Discourse Interface* conference (*Interface Discours-Prosodie, IDP*), held in Leuven (Belgium), from 11 to 13 September 2013. This conference is the fifth edition in a series of biannual conferences previously held in Aix-en-Provence (2005), Geneva (2007), Paris (2009), and Salford (2011). The general topic of these conferences is the relation between prosody and discourse. Research focuses on the way in which certain prosodic phenomena interact with discourse events and discourse structure, and vice versa. When examining the topics covered in these conferences, the following observations can be made:

- With respect to prosody, most papers focus on cues related to *pitch* (basically pitch movements, pitch contours). *Temporal aspects* (rhythm, pauses, speech rate, etc.) have received much less attention. This trend changes in the 2013 edition of IDP, where temporal structure is mentioned explicitly as a special topic in the call for papers.
- Various approaches and *frameworks to discourse analysis* are used in the papers, including the “discourse structure paradigm” (e.g. Geneva School, RST), conversation analysis and the interactionist approach.
- The *most common topics* covered in IDP conferences are: (a) prosody and information structure (in particular focus marking, in various languages), (b) the prosody-syntax interface (although not necessarily within the context of discourse analysis), (c) prosodic units and their relation to discourse units or syntactic units.
- More recent editions of IDP show an increase of studies on the prosodic cues of *speaking styles*, either individual or professional styles, and to a lesser degree regional variation of prosody.
- Less common research topics include: (a) prosody of reported speech, (b) the prosodic features of discourse markers, (c) deixis and anaphora resolution.
- Finally, some studies deal with more specific or more technical topics, which are frequently covered in specialized conferences: prosody in language acquisition, prosodic interferences in second language, prosody in speech synthesis, prosody and disfluency, etc.

The overall perspective of prosody in discourse entails the use of large-scale corpora of spontaneous, continuous speech, rather than read speech or other types of controlled speech. Over the editions of the conference, the relationship between prosody and discourse has been viewed from the perspective of phonetics, phonology, syntax, semantics, pragmatics, language acquisition, language processing, language pathology, stylistics and speech synthesis. Moreover, research was carried out in a variety of theoretical paradigms.

---

<sup>1</sup> Editor's contact: [Piet.Mertens@arts.kuleuven.be](mailto:Piet.Mertens@arts.kuleuven.be) , [anne-catherine.simon@uclouvain.be](mailto:anne-catherine.simon@uclouvain.be)

The call for papers of the 2013 edition of Prosody-Discourse Interface suggested five themes. The first three are about general features of prosody, whereas the other two are related to prosodic variation.

- **Prosodic boundaries**  
Theories about prosody often posit prosodic boundaries, which are supposed to play a central role in information structuring. But how does the listener identify such boundaries? Are they marked by dedicated prosodic forms (tones, contours)? If so, what are the acoustic and perceptual correlates of these forms? Or are prosodic boundaries only identified as such when they occur at particular positions in syntactic or discourse structure? If so, which are these positions? What are the functions of prosodic boundaries in discourse? Are boundaries marked in a language-specific way, or do they have language-universal features?
- **Pitch range, global and local**  
Does pitch range (the span between lower and higher pitches for a given speaker) remain constant in discourse, or does it change in such a way that stretches with narrow or wide register (local pitch range) may be identified? How does one detect register? What is the role of register (or register change) in discourse? How does the existence of register affect the identification of prosodic forms (such as tones)?
- **Speech rate and rhythm**  
Speech rate is commonly defined as the number of syllables per second. Still, it is clear that the duration of successive syllables varies considerably and that local speech rate changes dynamically. What could be an appropriate measure to specify local speech rate? What is the function of speech rate or change of speech rate in discourse? Is it affected by accommodation between speakers?
- **Regional variants of prosody**  
It is often assumed that a prosodic contour with a given function may differ from one regional variety of a language to the next. If so, how does one characterize these alternative realisations and how may they be detected in continuous speech?
- **Prosody and individual speaking styles**  
Past research has shown the existence of genres and even individual speech styles, with respect to prosodic properties. Do listeners identify such attributes and do speakers use them in a controlled way?

As expected, the papers presented at the conference covered a much broader range of topics. Among the more frequent ones were prosodic boundaries, prosody and discourse structure, and temporal aspects. Surprisingly, the topic of pitch range was represented by just one paper.

At the conference, two key-note speakers presented recent research on the pitch and timing phenomena in conversational interaction, a dynamic and joint activity where all speakers participate in the establishment of social relationships. The first key-note speech, “Timing and Prominence in Conversational Interaction”, by Petra Wagner (University of Bielefeld) focuses on speaker-adaptation or “entrainment”, and presents a formal model of the entrainment process, which relies on the identification of prominence. The second invited lecture, “Automatic analysis of speech prosody dynamics in social interactions: challenges and applications”, by Céline De Looze (Trinity College Dublin), presents general methods (in particular cluster analysis) and tools for the automatic analysis of prosodic aspects of conversational interaction.

For this special issue of JoSS, a selection of papers presented at the conference was made by the conference organizers, using the following criteria: (1) the study matches the general topic of prosody and discourse; (2) it matches one of the suggested topics of the 2013 conference; (3) the study uses large scale corpora of continuous speech; (4) it proposes a general approach, which may be applied to other data and to other languages; (5) it uses a novel approach.

Seven presentations were selected, the authors of which were invited to submit an extended version of their conference paper. This resulted in four submissions, three of which were accepted for publication. A brief presentation of these papers follows. Whereas the first article studies the automatic identification of pitch contours and their relation with syntactic structure, the second paper studies the prosody of particular syntactic constituents in the left periphery of the utterance, and the last paper deals with the prosodic properties of speaking styles in large-scale corpora.

The article by Philippe Martin, entitled “Analyse automatique de la structure prosodique d’*énoncés de styles variés*”, describes a procedure for the detection of pitch contour labels (C0, C1, C2, Cn) in French, following the intonation model proposed by the author. These pitch contours coincide with prosodic boundaries at the right end of an intonation unit. For syllables marked as prominent in the corpus annotation of C-Prom, the pitch movement within the vowel is evaluated relative to the perceptual threshold for pitch change (i.e. the glissando threshold). The nature of the contour depends upon the presence or absence of a glissando and on the (rising or falling) slope of that pitch movement. The proposed system is evaluated on the recordings of the C-Prom speech corpus, which includes seven speech styles. The resulting sequences of contour labels appear to be compatible with the intonation grammar in almost all cases, suggesting the validity of this grammar. The obtained results also illustrate the distribution of pitch contours and their sequences, depending on speech style. Since prominent syllables (or vowels) are identified using the corpus annotation, rather than from the speech signal itself, automatic detection of prominence should be added in order to obtain a fully automatic procedure.

The second paper, “*Les constituants introduits par des marqueurs de thématization en français face aux compléments spatio-temporels antéposés : une analyse prosodique*”, by Tom Velghe, studies the prosodic properties of sentence-initial constituents: either adverbials indicating space or time, or constituents introduced by *thematic markers* (TM) indicating a topic, such as “*en ce qui concerne*” and “*du point de vue de*”. It verifies whether the TM-constituents end in a major prosodic boundary, in the same way as does a left detached element (dislocation), the left part of clefts and pseudo-clefts, according to Mertens (2006). The corpus data shows this is indeed the case for 73% of the TM-constituents, whereas the major boundary is observed in only 43% of the adverbials. This difference is explained in terms of the semantic and syntactic properties of TM-constituents.

In their paper “*Prosodic features of situational variation across nine speaking styles in French*”, Tea Pršir, Jean-Philippe Goldman, and Antoine Auchlin study prosodic variation in French in various speech styles, using a general methodology which may be applied to other languages. They extend their earlier analyses in two ways: by exploring a more comprehensive set of speech styles and by the use of a much larger corpus (C-PhonoGenre, 8 hours). Speech recordings are characterized using four situational features: the presence of the audience at the time of speech, the intended broadcasting of the speech material, the degree of preparation of the speaker, the degree of interaction (monologue vs. interaction). Each situational feature accepts three alternatives, e.g. spontaneous vs. semi-prepared vs. prepared

speech. The main goal of the study is to investigate the correlation between situational features and acoustic or quantitative measures related to prosody. A large number (64) of measures is considered; these are related to pitch and temporal organisation of speech. Global measures include the *articulation ratio* (the proportion of articulated speech relative to the overall duration of speech including silent pauses) and the relative pitch variation (standard deviation divided by mean, for pitch values expressed in ST). Local measures include the proportion of final prominent syllables, the proportion of initial prominent syllables, the relative duration of the initial and final syllables in the stress group, segment durations, and so on.

The conference organizers would like to thank the editor-in-chief of the *Journal of Speech Sciences* for accepting to publish this selection of papers from the *Prosody-Discourse Interface* conference.

Campinas, 28th of July, 2014