

# Introduction

Sónia Frota, Gorka Elordieta, and Pilar Prieto

As the title indicates, *Prosodic Categories: Production, Perception and Comprehension* addresses the central question of the role played by prosody in language grammar and language processing. The eleven chapters of this book were developed from presentations to the *Third Tone and Intonation in Europe Conference (TIE3)*, hosted by the Universidade de Lisboa, Portugal, in September 2008, and all of them deal with different aspects of the definition, implementation and processing of prosodic categories. They present novel contributions to the understanding of key issues in prosodic theory, such as prosodic phrasing in production and comprehension, the relationship between intonation and pragmatics in speech production, speech perception and comprehension, the development of prosodic categories that convey specific pragmatic meanings, the characterization of the prosody of sentence modality, the role of pitch in quantity-based sound systems, the phonology of consonant conditioned tone depression across languages, and the encoding of intonational contrasts both in intonational and in tonal languages.

Exploring the intersection of phonology, phonetics and psycholinguistics, most of the chapters draw on empirical approaches to prosodic patterns in language: in particular, production, perception and comprehension experiments which include the prepared speech paradigm, the on-line speech production paradigm, conversational style and picture-naming production tasks, eye-tracking experiments using the real-world object manipulation paradigm, identification, discrimination and semantic scaling tasks, as well as perceptual experiments resorting to the gating paradigm. The production, perception and comprehension of prosodic categories is discussed in a wide array of languages (Swedish, Norwegian, Dutch, English, Bari Italian, Neapolitan Italian, Bengali, Estonian, Korean, Shanghai Chinese, Zulu, Shekgalagari, and Nle?kepmxcin), some of them underrepresented in the literature and others

---

S. Frota (✉)

Laboratório de Fonética & Lisbon Baby Lab (CLUL/FLUL), Universidade de Lisboa,  
Lisboa, Portugal  
e-mail: sfrota@fl.ul.pt

described for the first time with regard to these topics (like *Nie?kepmxcin* and *Shekgalagari*).

Aditi Lahiri and Linda Wheeldon's chapter investigates the prosodic grouping of phonological clitics and compound words, based on synchronic and diachronic data, as well as psycholinguistic evidence. They ask two key questions: do languages exhibit a preferred grouping of lexical words into larger prosodic constituents based on rhythmic principles? If so, does this prosodic grouping play a role in language production planning, that is in the processing involved in planning to produce speech? The authors entertain the hypothesis that, at least in a subset of languages, the natural grouping is trochaic, namely function words cliticize leftwards and compound words are left-headed. They set out to convincingly show that in Swedish, Norwegian, Dutch, English and Bengali leftwards attachment is indeed the natural prosodic grouping. Evidence comes from language change data showing the encliticization of the definite article in Scandinavian and the encliticization of auxiliaries in Germanic and Bengali. Additional evidence is provided by encliticization in Dutch, and unstressed words in English are also argued to show left attachment as their default pattern of prosodic phrasing. In short, trochaic grouping is claimed to be the preferred pattern for this set of languages, thus yielding phonological words that include clitics before their right edge and compound-like units that group together two phonological words.

The authors further show that research on the planning and articulation of speech allows the nature and size of the units the speaker uses in speech planning to be determined. They present experimental evidence, both from the prepared speech paradigm and the on-line speech production paradigm, that strongly suggests that the relevant unit for planning is the phonological word, and that the direction of attachment during cliticization (in Dutch) is leftwards. Specifically, on the basis of prepared speech production studies, they concluded that speech onset latency was a function of the number of phonological words in the utterance, and not of the number of lexical items, and that two word compounds are treated as one phonological word. The on-line speech production studies provided a clear indication that encliticization is the resultant prosodic grouping, as onset latencies to sentences that start with a word plus a clitic were slower than to sentences without a clitic. Taken together, the findings reported strongly support the hypothesis of trochaic grouping. However, as Lahiri and Wheeldon remark, this form of grouping is not universal: crucially, Romance languages have been shown to be by and large proclitic, favoring rightwards attachment of clitics and exhibiting compound-like structures which are right-headed (PeperKamp 1997; Vigário 2003; Hualde 2006/2007). Moreover, whereas Germanic languages extend the unstressed stretch on the right-hand side of phonological words by attracting unstressed elements, Romance languages tend to lose segments or syllables after the word stress (either by deletion or resyllabification – Harris 1988; Vigário 2003; Wheeler 2005). The reasons behind the difference in preferred prosodic grouping across languages (and related phenomena) constitute an important area of research

with implications for different fields, such as language typology, language variation and change and language acquisition.

Examining prosodic groupings at a higher level, Foltz, Maday and Ito's chapter is devoted to the production and processing of sentences with structural ambiguities. In their paper, these authors investigate the patterns of prosodic phrasing and semantic interpretation of sentences with two nouns in an NP followed by a relative clause, where the relative clause is sufficiently ambiguous to have either N1 or N2 as antecedents of the empty subject. An example of a sentence of this type would be *The brother of the bridegroom who swims was last seen on Friday night*. The investigation aimed at discovering the role of constituent length in determining prosodic boundaries as well as their strength. By varying the length of N1 (*the brother*, in the sample sentence above) and the relative clause, Foltz, Maday and Ito want to test the hypothesis that long constituents tend to be set off in their own prosodic phrases more often than short constituents (Carlson et al. 2001; Clifton et al. 2002; Fodor 1998, 2002; Watson and Gibson 2005). The other objective of Foltz, Maday and Ito is to see whether a full semantic processing and interpretation of an entire sentence affects prosodic boundary placement in a different way from a partial or on-line semantic processing. To this end, the authors carried out two production and perception experiments. In Experiment 1, subjects had to read and pronounce the target sentences (with filler sentences) on the fly, as they appeared on the screen, and then respond to a comprehension questionnaire by choosing between N1 and N2 as the antecedent of the subject of the relative clause (i.e., choosing between low and high attachment of the relative clause). In Experiment 2, the comprehension task was carried out first, and then the production task, after the subjects had read and processed the sentences semantically and assigned an antecedent for the subject in the relative clause. A comparison of the results of both experiments would be telling as to whether constituent length modulates prosody in all reading circumstances or only when the global sentential structure and its meanings are fully established by the reader/speaker.

The results for Experiment 1 show that the length of N1 determines prosodic boundary insertion, in the sense that the longer N1 is the greater the likelihood for prosodic boundary placement after it, but the results also show that relative clause length did not affect prosodic boundary placement and strength. Thus, these results do not support the hypothesis that long constituents are more likely to be segmented in an independent prosodic phrase (ip or IP). In the sentence comprehension task following the production task, subjects did not match their choices for antecedents with their production patterns, a result which is surprising in the light of previous work that had reported a closer match between prosodic boundary location and strength and attachment preferences (Carlson et al. 2001; Clifton et al. 2002). Experiment 2 shows that constituent length does not influence attachment preference as strongly as previously assumed, and that the preferred attachment option (low attachment) does not elicit stronger boundaries after N1 as often as expected. Unlike Experiment 1, however, Experiment 2 suggests that the length of the relative

clause combines with the length of N1 to affect the placement and strength of prosodic boundaries. That is, the results of Experiment 2 show that a more global consideration of the whole sentence is taking place when subjects can take time to process and comprehend the sentence before reading it aloud. Foltz, Maday and Ito explain the differences between the results of the two experiments in terms of the limits set by the eye-voice span of reading on the fly (as in Experiment 1), which does not allow calculating the length of the relative clause, opposed to the more generous time a subject has when reading a sentence without the pressure to utter it immediately (as in Experiment 2). In Experiment 1, the subjects would be guided by the presence of the syntactic boundary corresponding to the relative clause, more than by its length. In Experiment 2, on the other hand, the global knowledge of the syntactic and semantic structure of the sentence seemed to allow for a better rhythmic adjustment of the utterance.

The comparison between the two contexts for sentence production (reading on-line and after sentence comprehension) established by Foltz, Maday and Ito constitutes a novelty in studies of prosodic boundary insertion and their (mis) match with attachment preferences. The results and main conclusions of this work help advance our knowledge of the matching between sentence processing or comprehension and sentence production, as regards structurally ambiguous sentences. In particular, the results raise interesting questions as to how silent reading operates and how silent reading may reveal a lack of uniformity in implicit prosody that was not assumed before (cf. Fodor 1998, 2002; Jun 2003).

The next four chapters are all concerned with various aspects of the prosody of focus in four different languages. Kiwako Ito and Shari R. Speer's chapter has the goal of investigating the effect of the presence of a L + H\* pitch accent on pre-nominal adjectives in English. Even though it is well-known that the presence of a salient/focal pitch accent has the role of singling out a specific item from a larger set, what is less known is whether the semantics of the accented adjectives influences the activation of this contrastive interpretation. With this aim in mind, Ito and Speer conducted a pair of eye-tracking experiments designed to compare the effect of the presence of a prominent pitch accent on pre-nominal intersective color adjectives (Color experiment) and subjective size adjectives (Size experiment). Importantly, since subjective adjectives such as *big* or *old* require a relative interpretation, the hypothesis is that these types of adjectives may automatically evoke a notion of contrast (e.g., use of 'big' in 'Give me a big cup', which implies the presence of smaller cups). Contrastive interpretation was tested using the well-known real-world object manipulation paradigm, where participants followed pre-recorded instructions to decorate holiday trees. Eye fixation results revealed that in both experiments (the Color and the Size experiments) the presence of a prominent pitch accent (L + H\*) on the adjective facilitated the more rapid detection of the contrastive target (e.g., Hang a red/medium star. - Next, hang a *YELLOW/LARGE* star.). When L + H\* was infelicitously used in non-contrastive sequences (e.g., Hang a red/medium tree. Next, hang a *YELLOW/LARGE* ball.), results indicated no

inherent semantic advantage for the subjective adjectives, given that fixation times were not longer in the non-focal accent condition with the size adjectives than with the color adjectives. Moreover, contrary to the authors' prediction, the likelihood to fixate on the contrastive competitor was generally much higher for the Color than for the Size experiment. The results thus demonstrate that the presence of a focal prominence is evaluated online against the discourse context. When L + H\* is used felicitously on the adjective, it facilitates the detection of the target, regardless of the semantics of the adjective. By contrast, if L + H\* is used in an infelicitous way, it results in a slower detection of the correct target rather than as an increase in the fixations to the incorrect contrastive referents.

Importantly, while Ito & Speer's results confirm that the presence of a focal L + H\* pitch accent on pre-nominal modifiers activates a contrastive interpretation, they also found that the online comprehension of contrastive meanings is modulated by the discourse and visual context. Crucially, the unexpected bias toward a contrastive interpretation in the Color experiment uncovered a striking effect of the visual and discourse context in utterance interpretation rather than an inherent difference in adjective semantics. The authors convincingly attribute these differences to the salience/ease of the visual contrast that characterizes the display of the Color experiment as compared to the one of the Size experiment. One important methodological lesson to be learnt from these results, as the authors rightfully point out, is that experimental paradigms need to be controlled for easiness in referential comprehension within the visual field.

Aoju Chen's chapter is concerned with the phonological focus-marking in child language in Dutch. The information structure category focus, defined as the constituent that expresses new information in a sentence, exhibits specific phonological cues that are essential to focus-marking in adult Dutch: namely, information structure constrains both the placement of pitch accents and pitch accent choices in focused and unfocused positions. Children's ability to use phonological cues to mark non-contrastive narrow focus early on is inspected, in spontaneous (two-year-olds) and elicited speech (four- to eight-year-olds). On the methodological side, Chen's study puts forward an ingenious experimental set-up for data elicitation, that allows the collection of strictly comparable SVO declaratives with focused and unfocused NPs both in sentence-initial and sentence-final position. This is achieved by means of a picture-matching game where the production of trigger utterances is controlled for intonation pattern. The intonational analysis of child language was based on the Transcription of Dutch Intonation system (ToDI), which although designed to account for adult intonation has proven successful in dealing with child speech. The results have shown that children, two-year-olds included, use phonological means to mark focus, but not in an adult-like way. The use of accent placement and accent type to mark focus is acquired in a gradual fashion: two-year-olds do not use accent placement, but make a difference between non-downstepped accent types, used in focused words, and downstepped accents, used in unfocused contexts; four- to five-year-olds are already adult-like in using accent placement to realise the focused/unfocused contrast, but, unlike adults, show

no preference for the H\*L accent over the other accent types for focus in sentence-final position; seven- to eight-year-olds are largely adult-like in the phonological marking of focus.

The developmental path to phonological focus-marking proposed thus suggests a first step where focused/unfocused is generally equated to phonetically strong/weak, followed by the acquisition of the relationship between accent placement (accented/deaccented) and information structure, and finally of the relationship between accent type (first in sentence-initial position and later in sentence-final position) and information structure. As noted by the author, further analysis on younger children's utterances with final focus is called for to verify whether the phonetic strength distinction holds in this condition. Moreover, we would add, the prosodic phrasing of two-word utterances at this early stage needs to be carefully considered, as it may be the case that accent placement is not used to mark focus because each word forms its own phrase (as in single word utterances, or in successive single word utterances – Hallé et al. 1991; Behrens and Gut 2005; Frota 2010b). The proposed developmental path raises interesting questions for languages unlike Dutch, such as French (where pitch accent shape seems to play no role in focus marking), many other Romance languages (where pitch accent type matters but the Germanic accented/deaccented contrast does not exist in the intonation grammar), or Mandarin Chinese (where focus is marked by pitch range variation and duration). The effects of language-specific input on the path of development are a challenging topic for future cross-linguistic research on the interaction between prosodic categories and information structure in acquisition.

The contribution by Karsten Koch also deals with the interaction of intonation and information structure. Specifically, Koch provides a detailed phonetic analysis of prominence cues to focus and given information in *Nleʔkepmxcin* (Thompson River Salish). This is the first such study in any Salish language, thus providing new data as well as relevant insights from a yet largely unstudied endangered language. Based on informal and impressionist observations from the literature, the author sets out to test the hypothesis that *Nleʔkepmxcin*, although a stress language, does not mark the focus/givenness distinction by means of pitch accents or any kind of additional prominence/reduction of prominence in comparison with neutral, wide focus. The focus types under analysis are subject and object narrow focus, both non-contrastive and contrastive, and wide focus or focus-neutral. The examination of the most common acoustic prosodic correlates of focus and givenness, as reported in the literature on stress languages – namely, peak height, peak timing, local pitch range, peak intensity and accented vowel duration –, shows the absence of pitch cues in the marking of focus and givenness (the results of intensity and duration being inconclusive). This result, as argued by the author, makes Thompson Salish a typologically unusual stress-accent language, and has implications for the putative universality of constraints like STRESS-FOCUS and DESTRESS-GIVEN (e.g. Féry and Samek-Lodovici 2006), within stress languages.

The study undertaken by Koch, similarly to the work by Chen, points to the importance of cross-linguistic experimental research to establish the prosodic categories involved in the marking of information structure. In the case of Koch's study, the absence of acoustic pitch cues to prominence that are common in many stress languages is undoubtedly a relevant finding, although a phonological analysis of the Thompson Salish intonation system is not within the scope of this study. Such a finding also highlights the need for a global grammatical approach to focus marking. Thompson Salish, similarly to other languages where prominence cues are not used for focus marking, such as Wollof (Rialland and Robert 2001), marks focus by morpho-syntactic means, and thus it could be argued that the use of stress and accent features is analogous in grammatical function to the use of morphological and syntactic focus marking (Frota 2000, 2002). Under this view, prosodic categories need not be universally involved in focus marking, and their role in cuing information structure may have been overestimated by the study of Indo-European stress languages, as Koch duly suggests.

The chapter contributed by Kyunghee Kim investigates the influence of a number of prosodic and non-prosodic factors on the alignment patterns of the peak of the Accentual Phrase (AP) in Korean. In this language, the AP is one of the main units of prosodic analysis and it is demarcated by the tonal pattern THLH (T = L or H). Thus the AP contours can contain a peak which corresponds to an initial phonological H tone associated with the second syllable of an AP. The peak is phonetically realized either in the second or third syllables of the AP, and little is known about which are the factors that condition the alignment of this peak (see previous work on Korean intonation by Jun 2005). Two production experiments were conducted which consisted of casual style conversations with questions and elicited target answers. Experiment 1 aimed at investigating the potential effect of the number of phonological words in the AP, sentence length (presence or absence of a preceding AP), and focus type (narrow focus vs. broad focus) on peak placement. Results revealed that the realisation of narrow focus depends on AP length. In the short two-word AP, narrow focus is realised by earlier peak alignment. Number of phonological words in the AP has a significant effect on peak alignment, which is placed earlier in one-word APs than in two-word APs. Unexpectedly, the peak in the one-word AP was located systematically in the third syllable, as in two-word APs, indicating that the presence of a morpheme boundary is likely to affect the peak alignment. To test this hypothesis, Experiment 2 was carried out. The goal of this experiment was to test whether the alignment of the accentual peak was affected by the presence of an upcoming morpheme boundary and the presence/absence of semantic content in the following morpheme. The results indeed showed that accentual peak alignment is significantly affected by the presence of a morpheme boundary and by the semantic content of the following morpheme.

One of the most interesting results of the chapter by Kyunghee Kim is the fact that the alignment pattern of the accentual peak is systematically affected

by the location of a morpheme boundary. The location of the AP peak in Korean is confined to the AP initial morpheme, and the peak is aligned later as the morpheme becomes longer. Thus as the author argues, the H tone is associated with an edge of a morpheme. Yet this constraint is overridden by the semantic importance of the following morpheme. Even though previous work on tonal alignment patterns in European languages has highlighted the importance of prosodic factors on peak location, the results of this chapter demonstrate that in other languages the location of  $f_0$  peaks may directly depend on the presence of upcoming morpheme boundaries. Recent work by Prieto et al. (2010) has reported a similar phenomenon in two Romance languages, namely Spanish and Catalan. In these two languages, the presence of word boundaries also affects prenuclear peak location, and the target position of the peak can even be used by Catalan and Spanish speakers for word identification. These crosslinguistic findings stress the importance of carrying out typological work and encourage promising work in this area.

The following two chapters report on research on the perception of intonation categories and their relevance to meaning distinctions. Michelina Savino and Martine Grice conduct two perception experiments to investigate the role of pitch height variation in distinguishing between two different questions types in the Bari variety of Italian, namely yes-no information seeking question (Queries) and questions that challenge assumed given information (Objects). Previous work on the intonational marking of questions in Bari Italian has shown that pragmatic differences that are implemented in a gradient fashion may exhibit a discrete intonational marking, as in the case of the Query-Check dimension where different degrees of speaker's confidence are signalled by means of contrasting pitch accents ( $L + H^*$ ,  $H^* + L$ ,  $H + L^*$ ). The difference between Objects and the other questions types is also to be found in the accented syllable. Indeed, the nuclear pitch accent is the domain for the intonational marking of questions in Bari Italian (and not the boundary tone). However, Objects show what appears to be the same pitch accent as Queries ( $L + H^*$ ), but with a higher peak. The question thus arises whether the pragmatic distinction between Objects and Queries is signalled gradiently by intonation, or whether the peak height difference is phonological. Both an identification task and a discrimination task were carried out to examine whether the peak height difference is perceived categorically. Along the lines of other studies on pitch height differences in several languages (e.g. Chen 2003; Falé and Hub Faria 2006; Borràs-Comes et al. 2010), listeners responses in the semantically motivated identification task, together with the obtained pattern of Reaction Time measurements, clearly show that they are able to make a categorical interpretation of utterances as Query or Object on the basis of peak height variation only; however, and again in accord with many other studies, the results of the purely psychoacoustic discrimination task show that listeners are unable to discriminate between pairs of stimuli.

The results are interpreted in two relevant ways. First, the success obtained in the semantically motivated task, which is truly a linguistic task where subjects



must access linguistic knowledge on the categories available in the language and the way they are realized, points to the presence of two categories and therefore to the need to represent pitch height in the intonational phonology of Bari Italian. Second, the failure in the discrimination task, which is a psychoacoustic task that did not involve accessing linguistic knowledge, strongly suggests that this kind of task is not suitable to investigate intonational meaning contrasts, which have a semantic/pragmatic nature (see, as well, Chen 2003; Frota 2010a). The authors point to listener specific competence and acoustic memory restrictions as possible key factors affecting discrimination performance. In other studies, it has been argued that the problem may reside in the non-linguistic nature of the task, as semantically motivated discrimination tasks have been shown to provide different results (Schneider et al. 2009; Frota 2010a). The present chapter does offer an important contribution to the discussion about the approaches and methods to define prosodic categories. The chapter also questions the phonological nature of the intonational contrast investigated. In this study, as the authors rightfully mention, peak height is strictly equivalent to pitch range, and thus the results obtained do not directly inform the kind of phonological analysis that may best account for the pitch height categorical distinction, calling for future investigation.

The chapter written by Caterina Petrone and Mariapaola D'Imperio deals with the potential difference found between the prenuclear contours of Neapolitan Italian narrow focus statements and yes/no questions. As it is well-known from previous work by Mariapaola d'Imperio and colleagues (D'Imperio 2000; D'Imperio and House 1997), these two sentence types are distinguished by nuclear pitch accent alignment, namely early alignment (L + H\*) is found in narrow focus statements and late alignment (L\* + H) is found in yes/no questions. The chapter investigates two related questions: (a) the potential relevance of a tone which appears to be inserted at the right edge of the Accentual Phrase in the prenuclear contour (H in questions and L in statements), testing whether Neapolitan listeners are able to identify the contrast between questions and statements in the prenuclear region based on the edge tone difference; (b) the linguistic and paralinguistic meaning differences conveyed by this prenuclear edge tone. To study these questions, two experiments were run with gated stimuli, one with a forced-choice identification task and the other using a set of five semantic differential tasks. In the first experiment (Experiment 1), nine Neapolitan listeners heard three gates of the two sentence types for three separate sentences (with a control that contained the whole sentence) and they were asked to label each stimulus as either a question or a statement. If question/statement identification depended solely on the availability of nuclear accent information, listeners would not be capable to identify the target sentence type in early gates. Yet the results showed that the presence of prenuclear accent information was important: mean 'question' score for question base stimuli was above chance (67%), while it was around 37% for statement base stimuli. Moreover, when the prenuclear edge tone was present in the gate, question scores decreased for statement base stimuli (20%), suggesting that the

presence of this tone plays an important role in sentence type identification. The second experiment (Experiment 2) explored the potential contribution of the scaling of the target prenuclear edge tones L and H on the linguistic or paralinguistic meaning of this part of the contour. A set of five semantic differential tasks was run with five semantic scales that were selected on the basis of hypotheses about the linguistic and paralinguistic properties of the two accentual phrase tones, namely, 'commitment' (belief about the compatibility of the speaker's beliefs with those of the listener), 'potency' (or certainty or uncertainty about the content of his/her message), 'activity' (speakers' emotional involvement), 'evaluation' (speakers' sociability), and 'submission' (speakers' degree of authoritativeness). Nine Neapolitan listeners heard a set of gated utterances which contained a continuum in scaling between the L and the H prenuclear edge tones. The results revealed that only two out of the five scales varied with the scaling changes. Specifically, low edge values conveyed a higher degree of 'certainty', which progressively decreased as the edge height increased. Also, the mean involvement score progressively decreased as the height of the tone increased. The results of the two experiments pose a series of interesting questions to the study of prosody and meaning. While differences in the perception of the two modalities are demonstrated to be at work already in very early portions of the utterance and crucially at the point where the edge tone is available (see results of Experiment 1), differences in scaling of the edge tone significantly affect other listeners' semantic judgments.

This chapter represents a contribution to the still understudied topic of the relevance of prenuclear contours to meaning, as recent work on intonational phonology has mainly focused on the role of nuclear configurations. The results of Experiment 1 clearly challenge the idea that the nuclear configuration is the only cue that is relevant for the question-statement distinction. Tune meaning is clearly the result of the interaction between prenuclear and nuclear *f0* contours, and this issue calls for further investigation, as the authors point out. The results of Experiment 2 also provide with an interesting groundwork for starting the investigation of the relationship between the linguistic and paralinguistic semantic weight carried by prenuclear and nuclear contours in tune interpretation.

The chapter contributed by Pärtel Lippus, Karl Pajusalu and Jüri Allik deals with one of the central features of Estonian word prosody, namely, the three-way distinction in quantity (short or Q1, long or Q2, and overlong or Q3, which can be implemented by both vowels and consonants). While the primary cue for the three-way quantity distinction is the temporal pattern of the disyllabic foot, pitch patterns also seem to play an important role in the identification of the three-way quantity system. While in short and long categories the pitch falls at the end of the stressed syllable, in the overlong category the pitch falls in the first half of the stressed syllable. Previous research has shown that a conflicting combination of pitch and temporal cues can significantly affect target word identification. The main aim of the chapter is to investigate the role of pitch contour changes in the perception of the long vs. overlong quantity distinction. The authors performed

two identification experiments in which they manipulated different properties of the pitch patterns of a Q2 word without changing its duration. In Experiment 1, the locus of the fall was always in the middle of the accented syllable and the duration of the fall was varied in five steps. By contrast, in Experiment 2, the start of the fall was varied by five 20 ms increments and the pitch always fell during 50 ms (about 1/3 of the target vowel duration). In both experiments, listeners had to decide about the meaning of the words: 'send!' in case of Q2 and 'to get' in case of Q3. The results of Experiment 1 demonstrate that the duration of the fall is important for Q3 recognition, as the fall cannot be too short. The results of Experiment 2 show that the locus of the fall is crucial for Q3 perception. The pitch contour that triggers Q3 perception falls in the middle of the target vowel; yet if the pitch fall is too early or late during the vowel, Q2 is perceived. Pitch range is also found to be an important cue to quantity distinctions.

The chapter by Pärtel Lippus and collaborators represents an important contribution to our knowledge about the combined role of tonal and duration prosodic cues in the conveyance of lexical meaning in a language with a three-way quantity distinction. As the authors point out, the results of both experiments showed that several tonal features are essential for the perception of Q3, namely a significant locus of the falling *f0* pattern, pitch range of the fall and optimal length of the pitch movements. The potential interaction between these acoustic parameters is a challenging task for the future and for our understanding of the relationship between duration and *f0* parameters in the perception of speech.

The two last chapters in this volume discuss the phonology and phonetics of the uses of *f0* in tonal languages. Yiya Chen and Laura J. Downing establish a phonetic and phonological comparison of tone depressor consonants in two typologically unrelated languages, Shanghai Chinese and Zulu. Their objective is to rebate a previous proposal by Jessen and Roux (2002) that the depressor consonants in Xhosa (another Nguni language closely related to Zulu) and Shanghai Chinese can be characterized by the feature [slack voice], implemented phonetically the same way in the two languages, and that *f0* lowering occurs to compensate for absence of phonetic voicing of the depressor consonants. The data of Shanghai Chinese comes from previous work by Chen (2007), in which the effect of consonantal laryngeal features (aspirated, unaspirated and depressor) in different tone combinations across syllables was investigated. The data of Zulu comes from an experiment specifically designed by Chen and Downing to test the phonetic effect of depressor consonants, following Chen's (2007) methodology. The results of the experiment for Zulu show that the *f0* lowering effect of depressor consonants is different from that of depressor consonants in Shanghai Chinese; whereas in Zulu the *f0* level maintains low throughout the target syllable, in Shanghai Chinese it wanes much faster during the target syllable. Moreover, in Zulu *f0* lowering applies word-initially and word-medially, whereas in Shanghai Chinese there is less *f0* lowering word-medially. These results contradict Jessen and Roux's (2002) previous claim that depressor consonants behave the same way in the two languages. Chen and Downing

also show that Zulu implosives, which are fully voiced, do not lower *f* $\theta$  as depressor consonants do. This fact argues against a possible explanation of the *f* $\theta$  lowering effect of depressor consonants as a compensation for the loss of voicing of these consonants, as suggested by Jessen and Roux (2002). Chen and Downing propose an alternative account for the differences in *f* $\theta$  lowering between Shanghai Chinese and Zulu. Like Jessen and Roux (2002), they assume a feature [slack voice] for depressor consonants in the two languages, but they argue that the different phonetic effects this feature has in Shanghai Chinese and Zulu can be explained by phonological differences between the two languages. On the one hand, there are differences between Shanghai Chinese and Zulu regarding the domains of tonal specification: in Shanghai Chinese the domain is the syllable, whereas in Zulu it is the word. This would explain why in Shanghai Chinese the *f* $\theta$  lowering effect wanes faster than in Zulu. On the other hand, in Shanghai Chinese, the word-medial underlying tonal specifications of each syllable are lost by tonal sandhi processes, including the depressor register, so the different behaviour of depressor consonants word-medially could be explained as a way to compensate for the loss of tonal specifications, that is, as a way to maintain certain phonological specifications. In Zulu, there is no specific depressor register and there are no tonal sandhi processes word-medially that obliterate underlying tonal information, so there is no need to compensate or maintain anything.

Chen and Downing's account of the differences in the effects of *f* $\theta$  lowering between two tonal languages such as Shanghai Chinese and Zulu finds support in the idea that a single phonological feature can have different phonetic implementations in different languages (cf. Keating 1988; Kenstowicz 1994), and that these differences in phonetic implementation are governed or controlled by the phonology of each language (cf. Kingston and Diehl 1994). For the case at hand, Chen and Downing conclude that the phonological feature [slack voice] is present in Shanghai Chinese and Zulu but that the different phonetic realization of this feature is governed by (higher-order) phonological considerations. Hence, Chen and Downing's chapter contributes to our understanding of the phonological and phonetic properties of depressor consonants in two genetically unrelated tonal languages such as Shanghai Chinese and Zulu, refuting earlier proposals on the issue, as well as to our understanding of the phonology-phonetics interface in general.

In the final chapter, Larry M. Hyman and Kemmyonye C. Monaka address the different non-tonal strategies employed by tonal languages to convey grammatical meanings conveyed intonationally in non-tonal languages, such as sentence type. The central issue at stake is that tonal languages make use of *f* $\theta$  for establishing lexical and grammatical contrasts, so introducing additional tonal events at the phrasal or sentence level may give rise to conflicts with the tonal information already present at the word-level. As Hyman and Monaka point out, word-level tones may adopt three different types or degrees of receptiveness towards phrase- or utterance-level tones: accommodation or coexistence, submission or surrender, and avoidance or blockage. Shekgalagari,

a Bantu tonal language, represents the choice of accommodation, according to Hyman and Monaka. In this language, information of sentence type is signalled intonationally for certain sentence types, such as declarative sentences and citation forms, through the insertion of a L% boundary tone that is phonologically associated to the second mora of the lengthened penultimate syllable in phrase-final position. But for other sentence types, non-intonational cues are used. In ideophones, the sentence-final vowel is devoiced. In paused lists, the final syllable of each member of the list is lengthened. And in yes-no questions, wh-questions, imperatives, exhortatives, vocatives, exclamatives and monosyllables no segmental or suprasegmental cues are used. The marking of sentence types in Shekgalagari is interesting for a general theory of grammar, as it shows that the unmarked, general way of marking the majority of sentence types is through the absence of overt cues. That is, it could be concluded that in this language the unmarked cue of sentence type is the absence of intonation. Intimately related to this point, it is worth mentioning that another noteworthy aspect of the Shekgalagari system of marking sentence type is that it associates a phonologically marked cue (overt intonation and penultimate lengthening) to a pragmatically unmarked construction such as a declarative sentence.

Hyman and Monaka also show that when a sentence in Shekgalagari contains more than one of the above mentioned syntactic structures (e.g., when a wh-question ends in an ideophone), some cues override others, and thus the authors describe a hierarchy of sentence types or sentence-type cues. Interestingly, the phonologically unmarked sentence types (yes-no and wh-questions) dominate the marked sentence types (such as ideophones, lists and statements). Additionally, Hyman and Monaka reveal the existence of what they call emphatic declaratives, which are analyzed as abstract declaratives created out of the interrogative or exhortative sentences. Emphatic declaratives are marked by the same cues as regular declaratives or statements.

Hyman and Monaka end their chapter by raising the question of what exactly is intonation, and whether the non-tonal strategies observed in Shekgalagari to mark sentence types can be also considered intonational, given the fact that one of the main functions of intonation is to mark sentence types. This is a relevant question, from a theoretical point of view. If it is true that one of the main functions of intonation is to signal sentence type, it is also true that it is only one of several other functions, such as phrasing utterances into prosodic constituents, signalling focus, or expressing paralinguistic meaning. Hence an important issue for discussion in cross-linguistic research is whether it is appropriate to refer to the segmental marking of sentence types as *intonational*, or to restrict the term intonation to categories conveyed by suprasegmental cues.

The collection of studies presented in this volume will be of interest to a broad range of linguists and language researchers, such as phoneticians, phonologists, morphologists, syntacticians and semanticists interested in the syntax-phonology interface and the import of prosody to pragmatics and semantics. It will also be of interest to speech scientists, and to those with an interest on psycholinguistics and language acquisition and development.

We would like to thank all the contributors for their articles, and all the scholars who had agreed to review the chapters submitted to this book, as well as the helpful advice received from Springer's anonymous referees. Thanks are also due to Marisa Cruz for precious help in the preparation of the final book manuscript. This work has been partially funded by grants from Project FFI2009-07648/FILO and CONSOLIDER-INGENIO 2010 Programme CSD2007-00012 (both awarded by the Ministerio de Ciencia e Innovación, Spain), Project 2009 SGR 701 (awarded by the Generalitat de Catalunya, Spain), Project PTDC/CLE-LIN/108722/2008 (awarded by Fundação para a Ciência e a Tecnologia, Portugal), and by financial aid to the Research Group in Theoretical Linguistics HiTT (given by the Basque Government, Spain, ref. GIC07/144-IT-210-07) and to the Center of Linguistics of the University of Lisbon (given by Fundação para a Ciência e a Tecnologia, Portugal).

## References

- Behrens, H., and U. Gut. 2005. The relationship between prosodic and syntactic organization in early multiword speech. *Journal of Child Language* 32: 1–34.
- Borràs-Comes, Joan, Maria del Mar Vanrell Bosch, and Pilar Prieto. 2010. The role of pitch range in establishing intonational contrasts in Catalan. In *Proceedings of Speech Prosody 2010*, Chicago (<http://speechprosody2010.illinois.edu/papers/100103.pdf>).
- Carlson, Katy, Charles Clifton, Jr. and Lyn Frazier. 2001. Prosodic boundaries in adjunct attachment. *Journal of Memory and Language* 45: 58–81.
- Chen, Aoju. 2003. Reaction time as an indicator of discrete intonational contrast in English. In *Proceedings of Eurospeech 2003*, Geneva, 97–100.
- Chen, Yiya. 2007. The phonetics and phonology of consonant-F0 interaction in Shanghai Chinese. Talk present at the workshop on *Where Do Features Come From? Phonological Primitives in the Brain, the Mouth, and the Ear*. Paris, October 5<sup>th</sup>, 2007.
- Clifton, Charles Jr., Katy Carlson and Lyn Frazier. 2002. Informative prosodic boundaries. *Language and Speech* 45: 87–114.
- D'Imperio, Mariapaola. 2000. The Role of Perception in Defining Tonal Targets and their Alignment. Ph.D. Thesis, The Ohio State University.
- D'Imperio, Mariapaola, and David House. 1997. Perception of questions and statements in Neapolitan Italian. In *Proceedings of Eurospeech '97*, Rhodes, Greece, vol. 1, 251–254.
- Falé, Isabel, and Isabel Hub Faria. 2006. Categorical perception of intonational contrasts in European Portuguese. In *Proceedings of Speech Prosody 2006*, Dresden, 2–5 May 2006 (on CD-ROM).
- Féry, Caroline, and Vieri Samek-Lodovici. 2006. Focus projection and prosodic prominence in nested foci. *Language* 82(1): 131–150.
- Fodor, Janet Dean. 1998. Learning to parse? *Journal of Psycholinguistic Research* 27: 285–319.
- Fodor, Janet Dean. 2002. Psycholinguistics cannot escape prosody. In *Proceedings of Speech Prosody*, 83–88. Aix-en-Provence, France.
- Frota, Sónia. 2000. *Prosody and Focus in European Portuguese. Phonological Phrasing and Intonation*. New York: Garland Publishing.
- Frota, Sónia. 2002. The Prosody of Focus: a Case-Study with Cross-Linguistic Implications. In *Proceedings of Speech Prosody 2002*, Aix en Provence, 319–322.

- Frota, Sónia. 2010a. A focus intonational morpheme in European Portuguese: Production and Perception. In Gorka Elordieta and Pilar Prieto (eds.) *Prosody and Meaning*. Berlin: Mouton de Gruyter, submitted.
- Frota, Sónia. 2010b. Prosodic structure in early child speech: Evidence from intonation, tempo and coda production. Talk present at the Workshop on Prosodic Development. Universitat Pompeu Fabra, Barcelona ([http://www.fl.ul.pt/LaboratorioFonetica/texts/WPD\\_Frota2010.pdf](http://www.fl.ul.pt/LaboratorioFonetica/texts/WPD_Frota2010.pdf)).
- Hallé, P., B. Boysson-Bardies, and M. Vihman. 1991. Beginnings of prosodic organization: intonation and duration patterns of disyllables produced by Japanese and French infants. *Language and Speech* 34(4): 299–318.
- Harris, Martin. 1988. French. In Martin Harris and Nigel Vicent (eds.) *The Romance Languages*, 209–245. London: Routledge.
- Hualde, José Ignacio. 2006/2007. Stress removal and stress addition in Spanish. *Journal of Portuguese Linguistics* 5–2/6–1: 59–89.
- Jessen, Michael and Justus C. Roux. 2002. Voice quality differences associated with stops and clicks in Xhosa. *Journal of Phonetics* 30: 1–52.
- Jun, Sun-Ah. 2003. Prosodic phrasing and attachment preferences. *Journal of Psycholinguistic Research* 32: 219–249.
- Jun, Sun-Ah. 2005. Korean Intonational Phonology and Prosodic Transcription, Prosodic Typology. In Sun-Ah Jun (ed.) *Prosodic Typology. The Phonology of Intonation and Phrasing*, 201–229. Oxford: Oxford University Press.
- Keating, Patricia A. 1988. The phonology-phonetics interface. In Frederik J. Newmeyer (ed.) *Linguistics: The Cambridge Survey*, vol. I. *Linguistic Theory: Foundations*, 281–302. Cambridge: Cambridge University Press.
- Kenstowicz, Michael. 1994. *Phonology in Generative Grammar*. Cambridge, Mass.: Blackwell.
- Kingston, J. and R. L. Diehl. 1994. Phonetic knowledge. *Language* 70: 419–454.
- Peperkamp, Sharon. 1997. *Prosodic Words*. HIL Dissertations 34. The Hague: Holland Academic Graphics.
- Prieto, Pilar, Eva Estebas-Vilaplana, and Maria del Mar Vanrell Bosch. 2010. The relevance of prosodic structure in tonal articulation. Edge effects at the prosodic word level in Catalan and Spanish. *Journal of Phonetics*, 38/4: 688–707.
- Rialland, Annie, and Stéphane Robert. 2001. The intonational system of Wolof. *Linguistics* 39(5): 893–939.
- Schneider, Katrin, Grzegorz Dogil, and Bernd Mobius. 2009. German boundary tones show categorical perception and a perceptual magnet effect when presented in different contexts. In *Proceedings of Interspeech*, Brighton, 2519–2522.
- Vigário, Marina. 2003. *The Prosodic Word in European Portuguese*. Berlin/New York: Mouton de Gruyter.
- Watson, Duane and Edward Gibson. 2005. Intonational phrasing and constituency in language production and comprehension. *Studia Linguistica* 59: 279–300.
- Wheeler, Max. W. 2005. *The Phonology of Catalan*. Oxford: Oxford University Press.