Workshop on Prosody and Meaning
Barcelona | September 17-18, 2009

The Workshop on Prosody and Meaning is hosted by the Universitat Pompeu Fabra, the Universitat Autònoma de Barcelona and the Institut d’Estudis Catalans on September 17 - 18, 2009.

The goal of this conference is to bring together researchers working on the field of prosody and meaning. Recent developments in language research have increasingly put the spotlight on the phonological status of intonation and its relationship with meaning. This workshop is intended as a venue for exchanging ideas and methodologies and for stimulating discussions and collaborative work between researchers coming from different perspectives. The workshop consists of invited talks by members of the network and outside the network, a number of selected talks, and a poster session.

The Barcelona workshop is co-organized by Pilar Prieto (ICREA-UPF-UAB), Gorka Elordieta (EHU), and Joan Peytaví (U. Perpinyà IEC). The workshop is part of the activities of the research network Forms and Functions of Prosodic Structure (Carlos Gussenhoven, Yiya Chen -main convenors-, Gorka Elordieta, Sónia Frota, Aditi Lahiri, Pilar Prieto, Tomas Riad, Lisa Selkirk -coordinators).
### Thursday September 17

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.45–9.10</td>
<td>OPENING</td>
</tr>
<tr>
<td>9.10–10</td>
<td><em>Distinguishing contrastive, new and given information</em></td>
</tr>
<tr>
<td></td>
<td>Angelika Kratzer &amp; Lisa Selkirk (University of Massachusetts, Amherst)</td>
</tr>
<tr>
<td>10–10.50</td>
<td><em>A focus intonational morpheme in EP: production and perception</em></td>
</tr>
<tr>
<td></td>
<td>Sónia Frota (University of Lisbon)</td>
</tr>
<tr>
<td>10.50–11.20</td>
<td>COFFEE BREAK</td>
</tr>
<tr>
<td>11.20–12.10</td>
<td>Can intonational contours be lexicalised?</td>
</tr>
<tr>
<td></td>
<td>Sasha Calhoun (University of Edinburgh)</td>
</tr>
<tr>
<td>12.10–13</td>
<td><em>Prosodic Reflexes of the Interplay between Information Status and Focus-Background Structure in Spontaneous Speech</em></td>
</tr>
<tr>
<td></td>
<td>Stefan Baumann (Universität zu Köln)</td>
</tr>
<tr>
<td>13–15</td>
<td>LUNCH</td>
</tr>
<tr>
<td>15–15.50</td>
<td><em>Post-focus f0 suppression in Beijing Mandarin: now you see it, now you don’t</em></td>
</tr>
<tr>
<td></td>
<td>Yiya Chen (Leiden University)</td>
</tr>
<tr>
<td>15.50–16.20</td>
<td>Perceiving Focus Domains</td>
</tr>
<tr>
<td></td>
<td>Jason B. Bishop (UCLA)</td>
</tr>
<tr>
<td>16.20–16.50</td>
<td>Non-Exhaustive Answers to Wh-questions as Split Foci</td>
</tr>
<tr>
<td></td>
<td>Gorka Elordieta &amp; Aritz Irurtzun (Euskal Herriko Unibertsitatea)</td>
</tr>
<tr>
<td>16.50–17.50</td>
<td>COFFEE BREAK – POSTER SESSION 1</td>
</tr>
<tr>
<td>17.50–18.20</td>
<td>Syntax-Prosody Mapping and Topic-Comment Structure in Hungarian</td>
</tr>
<tr>
<td></td>
<td>Shinichiro Ishihara &amp; Balázs Surányi (Postdam University)</td>
</tr>
<tr>
<td>18.20–19.10</td>
<td>Phrasing, Focus and Meaning</td>
</tr>
<tr>
<td></td>
<td>Aditi Lahiri (University of Oxford)</td>
</tr>
</tbody>
</table>
Friday September 18

9–9.50  Prosody and context management in interactive discourse: a study of two different interaction scenarios
        Jill House (University College London)

9.50–10.40  Intonation and Pragmatic Effects
            Anne Wichmann (University of Central Lancashire)

10.40–11.20  COFFEE BREAK

11.20–12.10  Issues in investigating the mapping of prosody to meaning
              Laura Dilley (Bowling Green State University)

12.10–13  Local and global phrasing cues to Focus and Topic marking in Italian and French
          Mariapaola D’Imperio (Université d’Aix-en-Provence)

13–15  LUNCH

15–15.50  Meanings, shades of meanings and prototypes of intonational categories
          Barbara Gili-Fivela (Università del Salento – Lecce)

15.50–16.20  Syntactic disambiguation: the role of prosody
              Sónia Frota, Marina Vigário, Cátia Severino (Laboratório de Fonética - FLUL/CLUL)

16.20–16.50  The role of duration and tonal scaling as complementary cues in distinguishing presumptive from neutral yes-no questions
              Verònica Crespo-Sendra (U. Pompeu Fabra), Maria del Mar Vanrell (U. Pompeu Fabra), Pilar Prieto (ICREA-UPF)

16.50–17.50  COFFEE BREAK – POSTER SESSION

17.50–18.20  Prosody in German Sign Language
              Annika Herrmann (University of Frankfurt)

18.20–19.10  On linguistic and paralinguistic meanings of intonation
              Carlos Gussenhoven (Radboud University of Nijmegen)
POSTER SESSION 1

- **Factoring out Speaker Variation in Experimental Studies of Prosody: The Case of Association with Focus**
  M. Breen, M. Wagner, S. Shattuck-Hufnagel, E. Flemming, & E. Gibson (UMass Amherst, McGill University, MIT)

- **Non-local pitch range relationships in read and elicited speech**
  Alejna Brugos (Boston University)

- **On the prosodic marking of contrast in Romance sentence topic: evidence from Neapolitan Italian**
  Lisa Brunetti, Mariapaola D’Imperio & Francesco Cangemi (Laboratoire Parole et Langage, Aix-en-Provence)

- **Encoding focus in French: phrasing, deaccentuation, tonal patterns or all of them?**
  Aoju Chen (Max Planck Institute for Psycholinguistics) & Emilie Destruel (University of Texas)

- **The prosody of information structure in Paraguayan Guaraní**
  Cynthia G. Clopper & Judith Tonhauser (The Ohio State University)

- **Variations on Contrastive Topic Marking — Evidence from Mandarin Chinese**
  Noah Constant (University of Massachusetts, Amherst)

- **The prosody of ambiguous relative clauses in Spanish: a study of monolinguals and Basque-Spanish bilinguals**
  Irene de la Cruz-Pavía & Gorka Elordieta (Euskal Herriko Unibertsitatea)

- **Intonational encoding of topic and focus in adults with autism**
  Anne-Marie R. DePape (McMaster University), Aoju Chen (Max Planck Institute for Psycholinguistic), Geoffrey B. C. Hall (McMaster U.) & Laurel J. Trainor (McMaster U.)

- **Types of Topic in Turkish**
  Beste Kamali (Harvard University)

- **The accentuation of sentences with predicate composition**
  Manuela Korth (University of Stuttgart)

- **Three types of prosodic focus in Brazilian Portuguese: form and meaning**
  João Antônio de Moraes (Universidade Federal do Rio de Janeiro / CNPq)

- **The Prosodic Encoding of Information Structure in Beaver (Athabaskan)**
  Gabriele Müller (Westfälische Wilhelms-Universität Münster)

- **Focus, prosody and relevance in the Spanish of Buenos Aires**
  Leopoldo Omar Labastía & Alejandra Dabrowski (Universidad Nacional del Comahue, Argentina)

- **Relating prosody to pronominal reference: information structure versus switch-reference**
  James Śneed German (Laboratoire Parole et Langage, Aix-en-Provence)

- **The prosody of focus in Italian and German productions of Lecce (South Italy) speakers**
  Antonio Stella (CRIL - Università del Salento, Lecce)

- **Perceptual Robustness of the Tonal Center of Gravity for Contour Classification**
  Nanette Veilleux, Jonathan Barnes, Stefanie Shattuck-Hufnagel & Alejna Brugos (Simmons College, Boston University and Massachusetts Institute of Technology)
POSTER SESSION 2

- **The interplay of structural constraints and informational properties on PA selection**
  Giuliano Bocci & Cinzia Avesa (University of Siena & ISTC-CNR - Padova)

- **Early acquisition of form and meaning in Catalan and Spanish interrogatives**
  Joan Borràs-Comes (Parc Científic de Barcelona-UPF), Jill Thorson (Brown University), Verònica Crespo-Sendra (U. Pompeu Fabra), Maria del Mar Vanrell (U. Pompeu Fabra) & Pilar Prieto (ICREA-UPF)

- **Who weeded whom? German children’s use of prosodic cues in transitive constructions**
  Thomas Grünloh, Elena Lieven & Michael Tomasello (Max-Planck Institute for evolutionary Anthropology)

- **Prosodic expression of sentence-level pragmatic meaning in Akan**
  Frank Kügler & Susanne Genzel (Department of Linguistics & SFB 632 “Information Structure” Potsdam University)

- **Questions headed by the particle “que” in the Spanish spoken in Lleida: a prosodic bilingualism interference**
  Eugenio Martínez Celdrán, Ana M. Fernández Planas, Lourdes Romera Barrios & Josefina Carrera Sabaté (University of Barcelona, Laboratory of Phonetics)

- **The Effects of Prosody on the Interpretation of Novel Noun-Noun Combinations**
  Dermot Lynott & Louise Connell (School of Psychological Sciences, University of Manchester)

- **The effect of contextual interpretability on pitch entrainment in different age groups**
  Marie Nilsenová & Marc Swerts (CIW, Tilburg University, The Netherlands)

- **Conditions for tonal curl and stød in the Central Swedish variety of Eskilstuna**
  Tomas Riad (Stockholm University)

- **A quantitative implementation of the autosegmental-metrical model: the study of pitch and duration in Friulian**
  Paolo Roseano, Ana Maria Fernández Planas & Eugenio Martínez Celdrán (University of Barcelona, Laboratory of Phonetics)

- **Leading tone alignment in Occitan disapproval statements**
  Rafèu Sichel-Bazin (Universität Osnabrück-UPF)

- **Question intonation: for the layman and results on Brazilian Portuguese for the expert**
  Hubert Truckenbrodt (ZAS Berlin)

- **The role of pitch height in constraining the inferential space in Catalan yes-no questions**
  Maria del Mar Vanrell (U. Pompeu Fabra), Ignasi Mascarakó (Institut Menorquí d’Estudis), Francesc Torres-Tamarit (U. Autònoma de Barcelona) & Pilar Prieto (ICREA-UPF)

- **Prosodic Optionality or Syntactic/Semantic Choice?**
  Michael Wagner (McGill University)

- **Intonation in Discourse: Gradient or Categorical Behavior?**
  Margaret Zellers & Brechtje Post (Research Centre for English & Applied Linguistics, University of Cambridge)
Thursday September 17

**Oral (Invited and Selected) presentations**

_Distinguishing contrastive, new and given information_  
Angelika Kratzer & Lisa Selkirk (University of Massachusetts, Amherst)  

_A focus intonational morpheme in EP: production and perception_  
Sónia Frota (University of Lisbon)  

_Can intonation contours be lexicalised?_  
Sasha Calhoun (University of Edinburgh)  

_Prosodic Reflexes of the Interplay between Information Status and Focus-Background Structure in Spontaneous Speech_  
Stefan Baumann (Universität zu Köln)  

_Post-focus f0 suppression in Beijing Mandarin: now you see it, now you don’t_  
Yiya Chen (Leiden University)  

_Perceiving Focus Domains_  
Jason B. Bishop (University of California, Los Angeles)  

_Non-Exhaustive Answers to Wh-questions as Split Foci_  
Gorka Elordieta & Aritz Irurtzun (Euskal Herriko Unibertsitatea)  

_Syntax-Prosody Mapping and Topic-Comment Structure in Hungarian_  
Shinichiro Ishihara & Balázs Surányi (University of Potsdam, Hungarian Academy of Sciences)  

_Phrasing, Focus and Meaning_  
Aditi Lahiri (University of Oxford)
**Poster presentations**

**Factoring out Speaker Variation in Experimental Studies of Prosody: The Case of Association with Focus**
M. Breen, M. Wagner, S. Shattuck-Hufnagel, E. Flemming, & E. Gibson
(UMass Amherst, McGill University, MIT)

**Non-local pitch range relationships in read and elicited speech**
Alejna Brugos (Boston University)

**On the prosodic marking of contrast in Romance sentence topic: evidence from Neapolitan Italian**
Lisa Brunetti (Laboratoire Parole et Langage, Aix-en-Provence), Mariapaola D’Imperio (Aix-Marseille I/LPL, CNRS) & Francesco Cangemi (Laboratoire Parole et Langage, Aix-en-Provence)

**Encoding focus in French: phrasing, deaccentuation, tonal patterns or all of them?**
Aoju Chen (Max Planck Institute for Psycholinguistics) & Emilie Destruehl (University of Texas)

**The prosody of information structure in Paraguayan Guaraní**
Cynthia G. Clopper & Judith Tonhauser (The Ohio State University)

**Variations on Contrastive Topic Marking — Evidence from Mandarin Chinese**
Noah Constant (University of Massachusetts, Amherst)

**The prosody of ambiguous relative clauses in Spanish: a study of monolinguals and Basque-Spanish bilinguals**
Irene de la Cruz-Pavía & Gorka Elordieta (Euskal Herriko Unibertsitatea)

**Intonational encoding of topic and focus in adults with autism**
Anne-Marie R. DePape (McMaster University), Aoju Chen (Max Planck Institute for Psycholinguistics), Geoffrey B. C. Hall (McMaster U.) & Laurel J. Trainor (McMaster U.)

**Types of Topic in Turkish**
Beste Kamali (Harvard University)

**The accentuation of sentences with predicate composition**
Manuela Korth (University of Stuttgart)

**Three types of prosodic focus in Brazilian Portuguese: form and meaning**
João Antônio de Moraes (Universidade Federal do Rio de Janeiro / CNPq)

**The Prosodic Encoding of Information Structure in Beaver (Athabaskan)**
Gabriele Müller (Westfälische Wilhelms-Universität Münster)
Focus, prosody and relevance in the Spanish of Buenos Aires
Leopoldo Omar Labastía & Alejandra Dabrowski (Universidad Nacional del Comahue, Argentina)

Relating prosody to pronominal reference: information structure versus switch-reference
James Sneed German (Laboratoire Parole et Langage, Aix-en-Provence)

The prosody of focus in Italian and German productions of Lecce (South Italy) speakers
Antonio Stella (CRIL - Università del Salento, Lecce)

Perceptual Robustness of the Tonal Center of Gravity for Contour Classification
Nanette Veilleux, Jonathan Barnes, Stefanie Shattuck-Hufnagel & Alejna Brugos (Simmons College, Boston University and Massachusetts Institute of Technology)

Friday September 18

Oral (Invited and Selected) presentations

Intonation and Pragmatic Effects
Anne Wichmann (University of Central Lancashire)

Issues in investigating the mapping of prosody to meaning
Laura Dilley (Bowling Green State University)

Local and global phrasing cues to Focus and Topic marking in Italian and French
Mariapaola D’Imperio (Aix-Marseille I/LPL, CNRS)

Meanings, shades of meanings and prototypes of intonational categories
Barbara Gili-Fivelá (Università del Salento – Lecce)

On linguistic and paralinguistic meanings of intonation
Carlos Gussenhoven (Radboud University of Nijmegen)

Prosody and context management in interactive discourse: a study of two different interaction scenarios
Jill House (University College London)

Syntactic disambiguation: the role of prosody
Sónia Frota, Marina Vigário, Cátia Severino (Laboratório de Fonética - FLUL/CLUL)
The role of duration and tonal scaling as complementary cues in distinguishing presumptive from neutral yes-no questions
Verónica Crespo-Sendra (U. Pompeu Fabra), Maria del Mar Vanrell (U. Pompeu Fabra), Pilar Prieto (ICREA-UPF)

Prosody in German Sign Language
Annika Herrmann (University of Frankfurt/Main)

**Poster presentations**

The interplay of structural constraints and informational properties on PA selection
Giuliano Bocci & Cinzia Avesa (University of Siena & ISTC-CNR - Padova)

Early acquisition of form and meaning in Catalan and Spanish interrogatives
Joan Borràs-Comes (Parc Científic de Barcelona-UPF), Jill Thorson (Brown University), Verónica Crespo-Sendra (U. Pompeu Fabra), Maria del Mar Vanrell (U. Pompeu Fabra), Pilar Prieto (ICREA-UPF)

Who weeded whom? German children’s use of prosodic cues in transitive constructions
Thomas Grünloh, Elena Lieven & Michael Tomasello (Max-Planck Institute for evolutionary Anthropology)

Prosodic expression of sentence-level pragmatic meaning in Akan
Frank Kügler & Susanne Genzel (Department of Linguistics & SFB 632 “Information Structure” Potsdam University)

Questions headed by the particle “que” in the Spanish spoken in Lleida: a prosodic bilingualism interference
Eugenio Martínez Celdrán, Ana M. Fernández Planas, Lourdes Romera Barrios & Josefina Carrera Sabaté (University of Barcelona, Laboratory of Phonetics)

The Effects of Prosody on the Interpretation of Novel Noun-Noun Combinations
Dermot Lynott & Louise Connell (School of Psychological Sciences, University of Manchester)

The effect of contextual interpretability on pitch entrainment in different age groups
Marie Nilsenová & Marc Swerts (CIW, Tilburg University, The Netherlands)

Conditions for tonal curl and stød in the Central Swedish variety of Eskilstuna
Tomas Riad (Stockholm University)
A quantitative implementation of the autosegmental-metrical model: the study of pitch and duration in Friulian
Paolo Roseano, Ana María Fernández Planas & Eugenio Martínez Celdrán (University of Barcelona, Laboratory of Phonetics)

Leading tone alignment in Occitan disapproval statements
Rafèu Sichel-Bazin (Universität Osnabrück-UPF)

Question intonation: for the layman and results on Brazilian Portuguese for the expert
Hubert Truckenbrodt (ZAS Berlin)

The role of pitch height in constraining the inferential space in Catalan yes-no questions
Maria del Mar Vanrell (U. Pompeu Fabra), Ignasi Mascaró (Institut Menorquí d’Estudis), Francesc Torres-Tamarit (U. Autònoma de Barcelona), Pilar Prieto (ICREA-UPF)

Prosodic Optionality or Syntactic/Semantic Choice?
Michael Wagner (McGill University)

Intonation in Discourse: Gradient or Categorical Behavior?
Margaret Zellers & Brechtje Post (Research Centre for English & Applied Linguistics, University of Cambridge)
Abstracts
Thursday September 17
Oral (Invited and Selected) Presentations
The question of how linguistic theory should break down the dimension of “information structure” that includes contrastiveness, newness and givenness continues to be a subject of debate. This paper defends the three-way distinction between given, new, and focus of contrast originally proposed in Chafe 1976. This proposal echoes a recent proposal by Féry and Samek-Lodovici 2006 and developed in Selkirk 2007.

Evidence from patterns of prosodic prominence in English is presented that supports the hypothesis that the theory of grammar makes a representational distinction between contrastive focus and discourse-newness in the syntax and in its interface with phonology/phonetics. The phonological/phonetic evidence comes from an experimental investigation (Katz and Selkirk, 2006, in preparation) which compares the prosody of productions of all-new sentences with the prosody of sentences that combine putative contrastive focus constituents and constituents qualifying only as discourse-new. It turns out that while the distribution of pitch accents in these different sentence types is the same, the pattern of relative phonological/phonetic prominence in these sentences varies according to their composition in terms of contrastive and/or new constituents.

We will suggest in addition that a system which gives morphosyntactic representation to focus of contrast (FoC-marking) and to givenness (G-marking) but which leaves newness morphosyntactically unmarked has the right consequences for theories of the interfaces of syntax with sentence prosody on the one hand and with semantics on the other. On the phonology side, it is shown that all-new sentences receive a phonological interpretation that is based on general phonological and interface constraints that make no appeal to the information structure feature make-up of the sentence. On the semantics side, renditions of the Rooth 1992 theory of alternatives focus and the Schwarzschild 1999 theory of givenness are combined with a set of syntax/semantics interface constraints to provide an account of the interpretation and distribution of constituents which are FoC-marked, G-marked, and/or unmarked for either.

The typological predictions of our proposal are briefly explored: whether FoC-marking or G-marking are expressed in sentence prosody varies (independently) from one language to the next. Some languages show no prosodic reflexes of these morphosyntactic contrasts at all, instead defaulting to the types of unmarked sentence prosody found in all-new sentences.
Production studies of the intonational signalling of focus in EP have shown that focus is expressed by a specific pitch accent type, thus revealing a systematic contrast between nuclear accents associated with different meanings (Frota 2000, 2002a,b, Fernandes 2007). In the case of declaratives, the contrast is essentially realized as an alignment difference: H+L* (neutral accent) and H*+L (focus accent). The H*+L pattern is accompanied with focus-related lengthening. A pilot perceptual study using natural stimuli has shown that subjects are able to distinguish between members of neutral/focus minimal pairs and to match them to the appropriate production context (Frota 2000). The present research revisits the production contrast and investigates its categorical nature by means of 4 perception experiments which are all semantic tasks: (1) a context-matching identification task; (2) a context-matching discrimination task; (3) a semantic matching task; (4) a semantic scaling task. In experiments 3 and 4 reaction times are also obtained. Based on natural neutral (N) and focus (F) renditions, a N to F and a F to N continua were created by manipulating alignment of H and L targets. This design allows us to assess the relative contribution of alignment and lengthening to the neutral/focus contrast, as well as the gradual or categorical implementation of the distinction in perception. The present research is also relevant to recent discussion about the nature of intonational contrasts, and the various approaches and methods to define prosodic categories (Gussenhoven 2006, Schneider et al. 2006, Gili Fivela 2008, inter alia).
Can intonation contours be lexicalised?
Sasha Calhoun
University of Edinburgh
sasha.calhoun@ed.ac.uk

I present the first steps in a research program rethinking how intonation is represented and stored cognitively. My proposal is that intonation contours are stored in the first instance with the utterance they are produced on. The link between the contour and the words decays if it is not sufficiently discriminable and robust; however, if frequent enough, intonation contours can be lexicalised, i.e. associated with particular lexical items or discourse units. If true, this leads us to a very different way of conceptualising how intonational meanings, such as implicature, arise.

In recent years, evidence has been growing for exemplar-based models from the phoneme to the construction level (e.g. Pierrehumbert 2001, Bybee 2006). The idea is that linguistic units are stored as “clouds” of remembered instances. These exemplars include much more phonetic detail and are tagged for a much richer range of information about their occurrence, including semantic and pragmatic information and social context, than in standard models. Bybee (2006) proposes that, given this approach, pragmatic implications of specific utterances may be stored with an utterance if they are frequent enough. Those implicatures then spread through analogy to other similar utterances. This may even be the more common way for implicatures to arise, rather than the standard account that they are generated dynamically given a particular context.

To my knowledge, this approach has not been explored in explaining the “meanings” of intonation contours. However, it seems a promising way to explain why the very precise implicatures which arise from certain word-contour pairings (consider, for example, different intonational realisations of really?), do not seem to accord with the very broad “meanings” claimed for different accent and boundary tone types, e.g. “not mutually believed” or “speaker oriented” (e.g. Pierrehumbert & Hirschberg 1990, Steedman 2000). There is already some evidence for the role of frequency in accenting and intonation: Nenkova et al. (2007) have shown that the frequency with which a lexeme is accented is highly predictive of the likelihood that a token of that lexeme will be accented; Dainora (2002) found strong statistical associations between tonal accent type and following boundary type; while Walsh et al. (2008) have shown variation within intonation categories is strongly related to frequency.

In the first stage of testing this approach to deriving intonational meaning, I am carrying out a statistical analysis of the association between tonal pitch accent and boundary tone types, and lexical identity in a large corpus of conversational speech, Switchboard (Godfrey 1992). I am using a portion of the corpus, comprising 76 conversations (approximately 97,000 words), which has been manually annotated for accents and phrasing (Calhoun et al. to appear). From this I am using automatic pitch parameterisation techniques to extract “low-rise” contours (i.e. ToBI L* or L*+H followed by LH% or HH%). I have chosen these contours as they are associated with highly marked intonational meanings which are likely to give rise to the kinds of implicatures I am interested in (e.g. see Ward & Hirschberg 1985, Pierrehumbert & Hirschberg 1990). I will test whether there are lexical items or phrases which are more likely to be associated with these contours, than they are to be accented in general. If such statistical associations are found, this is evidence that frequency plays a role in the choice of intonation contour for an utterance, and therefore that intonation contours can be lexicalised. Assuming this study is successful, I plan, in future work, to test whether these associations affect lexical and intonational perception, and crucially the derivation of intonational meaning.
References


The talk reports on a preliminary investigation of the interplay between different aspects of information structure and their reflexes at various levels of linguistic description – in particular prosody – in spontaneous German monologues. The aspects of information structure we are particularly concerned with are levels of cognitive activation (or information status of discourse constituents) on the one hand and focus-background structure on the other, both of which we subsume under the cover term ‘informativeness’.

Following Lambrecht (1994), we assume that a referent’s information status has structural correlates – not only in written texts but also in spoken language, even if it is spontaneous and thus possibly ‘fragmentary’. These correlates may be found in (morpho-)syntax (word order, part-of-speech, definiteness, syntactic function), lexical semantics (sense relations between antecedents and anaphora) and phonology (in this case prosody: accentuation and phrasing). Focus-background structure also has formal, structural correlates. For instance, focus can be marked in syntax by non-default word order and/or by focus operators.

Although the marking of information status and focus both strongly depend on prosody for deriving the meaning of the sentence, one goal of the talk is to identify the textual cues to prominence which serve to mark informativeness. We refer to this as ‘Accumulated Prominence from Text’ (APT). This level of analysis and annotation takes place using the orthographic transcription without access to the speech signal, so as to ensure that the different levels are kept distinct. Working from the speech signal, we will discuss how far a constituent’s degree of cognitive activation and its role in the focus-background structure of an utterance are marked by a conglomerate of cues which, together, can be referred to as ‘Accumulated Prominence from the Speech Signal’ (APSS).

We shall ascertain the role of (i) categorical cues to prominence at the phonological level, expressed e.g. by the position and type of pitch accent (Baumann et al. 2006), and (ii) gradient cues at the phonetic level, such as accent peak height and timing, pitch excursion and duration of constituents of varying sizes (Baumann et al. 2006, 2007). Additionally, the degree of accentuation, or accent strength, is taken into account as well as the concepts of ‘secondary accents’ (e.g. Büring 2006) and ‘phrase accents’ (Grice et al. 2000), both of which have been proposed as markers of semi-active information and embedded focus (e.g. Halliday 1967).

By using a multi-layer annotation system comprising APT and APSS we propose a weighting procedure (inspired by studies from Rietveld & Gussenhoven 1995 and Wichmann et al. 2000 on pitch target alignment) to obtain a prominence value which in turn indicates the degree of informativeness of discourse constituents. This will help us to evaluate the role played by prosody and the degree to which it interacts with textual cues. The ultimate goal would be to predict the likelihood for a certain type and strength of accent to be used to encode a certain degree of informativeness – not only for laboratory speech but also for spontaneous speech.
References


In speech communication, the same string of words is often pronounced differently, depending on communicative contexts. Consider the following examples.

(1) Mary traveled in TIBET last year. (She did not travel in Hong Kong).
(2) Mary TRAVELED in Tibet last year. (She did not work there).

If the speaker intends to emphasize that the place Mary went to was Tibet, not Hong Kong, Tibet would be contrastively focused (indicated with capital letters) and is typically pronounced with prosodic prominence. As a contrast, the acoustic realization of Tibet in (2), as given information in a post-focus position, sounds much less prominent (or more reduced).

Much work has been done on how prosodic prominence is instantiated in different languages to package an utterance and integrate it into the information flow of ongoing discourse. In languages such as Standard Chinese, where F0 changes indicate lexical contrasts, it is often reported that focus is realized via pitch range manipulation (e.g., Jin 1996, Xu 1999). Specifically, “the pitch range of the focused region is expanded; that of the post-focus region compressed; and that of the pre-focused region left largely neutral” (Xu 2005: 235). Chen (2003) and Chen & Gussenhoven (2008), however, argue that focus does not just introduce pitch range manipulation. Rather, the effect of focus is better accounted for by appealing to an abstract notion of prosodic prominence. Specifically, it is proposed that a focused element in Standard Chinese is associated with high-level prosodic prominence of the utterance. Such structural prominence is manifested in the greater articulatory force that leads to more distinctive realization of tonal contours over the focused constituent.

In this study, we report data on post-focus tonal realization which argues further against mere manipulation of pitch range as a function of focus status. All four lexical tones in the post-focus condition were elicited from 5 speakers of Beijing Mandarin in different tonal contexts. Results show that while post-focus lexical tones may be realized with a compressed F0 range, in some tonal contexts post-focus lexical tones were realized with an F0 range that was expanded much more than their pre-focus counterparts (by comparison with data reported in Chen & Gussenhoven 2008).

The figure shows the F0 range of Rising and Falling tones in the post-focus condition, compared to the base-line pre-focus condition, uttered in two preceding contexts (Preceding tone High vs. Preceding tone Low). When the preceding tone was High (i.e., P-High), there was post-focus F0 range suppression in the Rising tone but an F0 range expansion in the Falling tone. When the preceding tone was Low (i.e. P-Low), both tones showed an effect of F0 range expansion, though with a much greater magnitude in the Rising tone. Usually, despite the sometimes similar F0 range across the focus conditions, what differentiates the post-focus from the pre-focus and on-focus lexical tones (reported in Chen & Gussenhoven 2008) was the degree of distinctiveness in their F0 contours (not shown in the abstract). We argue that the lack of distinctiveness in post-focus condition is due to the weak implementation of the lexical tones, as they are associated with prosodically non-prominent constituents. Such hypo-articulation also makes it possible for the preceding focused lexical tones to exert a strong carry-over influence on the post-focus tones, which sometimes results in post-focus pitch range expansion. Implications of these results on the cross-linguistic relation between prosody and information structure encoding will be discussed.
F0 realization of the Rising and Falling tones, preceded by High (P-High) or Low (P-Low) tone and followed by a Rising tone.
Perceiving Focus Domains
Jason B. Bishop
University of California, Los Angeles
j.bishop@ucla.edu

A recurrent finding in speech perception research is that a listener’s perception of the speech stream is an integrative process, an interpretation of the acoustic signal partially colored by non-signal-based factors. Frequently cited examples of linguistic knowledge guiding perception are of the phoneme restoration variety—listeners are shown to use lexical and semantic information (Warren 1970; Samuel 1981) or native language phonotactics (Dupoux et al. 1999) to perceive segmental properties absent in the acoustic signal. The vast majority of research in this vein, however, has been limited to the perception of vowels and consonants; much less is known about how expectations based in linguistic knowledge are reflected in the perception of prosodic information, such as phrasing and prominence.

One recent line of research, however, concerns itself with how a sentence’s information structure, in particular its focus structure, contributes to the perception of prosody. Vainio and Järvičkivi (2006), for example, demonstrate that judgments about the location of sentence-level prominence by Finnish-speaking listeners are influenced by focus structure such that words in a syntactic focus position are more likely to be perceived as prominent, even when the marked word order is not accompanied by any distinctive acoustic marking.

Whereas Vainio and Järvičkivi were interested in prominence perception mediated by a (word-order dependent) focus/neutral distinction, the present study investigates how prominence perception for pitch accents is influenced by the size of the focus domain containing it. Data were analyzed from an experiment in which linguistically untrained English-speaking listeners provided prominence ratings for verbs and nuclear-accented nouns for sentences in three different pragmatic contexts, each triggering a different focus domain: NP-focus, broad VP-focus or Sentence-focus (see Fig. 1). To test for the independent effect of focus domain size on prominence ratings, listeners heard question-answer pairs in which the test sentence (the answer) was held constant. That is, listeners always heard the same production of an answer sentence such as John bought a BOOK (in all cases, actually an answer to a VP-focus question), but following different questions. Results showed that listeners’ responses were sensitive to this information structural manipulation. Prominence ratings for nuclear accents (in relation to potential prenuclear accents) were inversely related to focus domain size, nuclear accents being judged more prominent as the focus domain narrowed (RM ANOVA [(2,16)F = 8.2, p = .001]), illustrated in Fig. 2. A post hoc pairwise comparison, however, showed the effect only significantly distinguished two categories: narrow NP-focus versus VP-focus/Sentence-focus.

Results are considered in light of studies suggesting focus domains can be distinguished (1) in speech production—both phonetically (see Löffstedt 2006 for English; Baumann 2007 for German), as well as phonologically (Jun, in progress, for English and Korean); and (2) by listeners’ judgments about appropriateness in context (Gussenhoven 1983; Welby 2003). The results of the experiment presented here support the findings of these studies, and extend them by showing that focus domains are not only realized and identifiable in the signal, but like other kinds of linguistic knowledge, are also projected onto the signal by listeners.
Figures

Q1: What happened?
A: John bought a book.
[---Sentence Focus---]

Q2: What did John do?
A: John bought a book.
[---VP Focus---]

Q3: What did John buy?
A: John bought a book.
[---NP Focus---]

Fig. 1. Question-answer pairs eliciting focus domains varying from broad to narrow. The same recording of the answer appeared in three contexts.

References


Usual analyses of the semantics of questions à la Hamblin-Karttunen propose that an answer to a wh-question (1b) picks up one proposition of the denotation of the question (1a) (which is the set of propositions obtained by the substitution of the wh-phrase by alternatives that match it in semantic type). This gives the question a complete answer. However, there are other cases where an answer to a wh-question may provide such a proposition but without answering the question fully. E.g., (1a) could be answered as in (1b) but implying that there might be other potential lovers of Paula, for which the speaker has no evidence; i.e., the open set denoted by the question is not cancelled by the answer. We refer to this type of answers as Non-Exhaustive Narrow Focus (NENF), opposed to Exhaustive Narrow Focus (ENF). We present evidence from French, Spanish and Basque showing that NENF is encoded intonationally. In subject wh-questions, in NENF the subject does not receive main prominence and it is the verb that receives it. Interestingly, these features are also observed in verum focus constructions, where the polarity is the focus of the sentence (cf. the dialogue in (2)).

We designed a production experiment with a series of answers to wh-questions focusing on the subject (cf. (3)-(5) as examples for French, Spanish and Basque). We instructed the native speakers to produce the answers as ENF and NENF, with the same word order. We also had verum focus constructions (VF) (cf. (6)-(8)), in order to compare their intonational patterns with those of NENF. Three native speakers of each language read eighteen sentences as answers to triggering questions, repeated twice (162 utterances in total). We measured the F0 maximum in the subject and in the verb, the difference in Hz between the two F0 maxima, and peak alignment with the accented syllable. In ENF, the subject has the most prominent accent in the utterance, and the verb has a much smaller peak. In NENF, the subject does not display main prominence, and in Spanish and French it ends in a continuation rise, with peak displacement in Spanish. The verb receives main prosodic prominence, with a much higher peak than in ENF. On the other hand, there were no consistent differences for subject and verb intonation between NENF and VF: the verb (i.e., the polarity) has the most prominent accent, with similar peak heights in VF and NENF, whereas the subject ends in a continuation rise (in Spanish and French and in one Basque speaker) and may show peak delay (in Spanish and in one Basque speaker). Thus, although in NENF the subject responds to a wh-variable, NENF is prosodically similar to VF and different from ENF.

Our analysis of the results is that in NENF, speakers introduce an additional focal feature on the polarity so as to convey the meaning that they cannot provide an exhaustive reference for the variable in the question. In ENF, main prosodic prominence on the constituent K that answers for the variable x conveys the meaning that the pairing of the alternative values raised by the wh-phrase in the question and the alternative values for the polarity (yes, no) is {K, yes}, and {all other alternatives, no}, cf. (9)). In NENF, speakers only commit to asserting that {K, yes}, but not {all other alternatives, no}. Thus, for a question like (1a), in a NENF answer a speaker only asserts that Mary loves Paula (i.e., the Subject-Polarity pair {Mary, yes}), but does not close other pairs (e.g., {John, yes/no}, {Peter, yes/no}) (cf. (10)). The value of the polarity of the possible alternatives is thus not resolved, and the polarity of the pair {Mary, yes} is highlighted. In fact, we make the claim that NENF constructions are split foci constructions, having both the subject and the polarity as focal, with the prosody we have summarized above.
Examples

(1) a. Who loves Paula?
   \{\text{love}(x, p) \mid x \in E\} = \{[[\text{Mary loves Paula}]], [[\text{John loves Paula}]], [[\text{Peter loves Paula}]], [[\text{Sarah loves Paula}]], [[\text{George loves Paula}]]\}

   \text{b. Mary loves Paula.}
   [[\text{love}(m, p)]] = [[\text{Mary loves Paula}]]

(2) A: - I’m not sure whether Mary loves Paula.
   B: - Mary does love Paula.

(3) A: Qui a amené le boudin?
   ‘Who brought the blood sausage?’
   B: Didier l’a amené / Didier a amené le boudin
   ‘Didier brought it / the blood sausage’

(4) A: ¿Quién devolvió el dinero?
   ‘Who returned the money?’
   B: Valerio lo devolvió / Valerio devolvió el dinero
   ‘Valerio returned it / the money’

(5) A: Nok eran dau ardaua?
   ‘Who has drunk the wine?’
   B: Nagorek eran dau ardaua
   ‘Nagore has drunk the wine.’

(6) A: Didier disait qu’il allait amener le boudin, mais je ne suis pas trop sûr qu’il l’ait fait...
   ‘Didier said that he would bring the blood sausage, but I’m not sure he actually did’
   B: Didier a amené le boudin.
   ‘Didier did bring the blood sausage’

(7) A: Valerio dijo que iba a devolver el dinero, pero no sé si lo ha hecho.
   ‘Valerio said he would return the money, but I don’t know whether he’s done it’
   B: Valerio devolvió el dinero.
   ‘Valerio did return the money’

(8) A: Nagorek ardaua erango ebala esan eban., baiña ez dakitx eran badau.
   ‘Nagore said she would bring wine, but I don’t know if she did’
   B: Nagorek eran dau ardaua.
   ‘Nagore has drunk wine’

(9) Subj. Pol.
    
    Mary  yes
    John  no
    Peter
    ……”

(10) Subj. Pol.
    
    Mary  yes
    John  no
    Peter
    …”
Background & Objectives: In this study we investigate the intonation of Hungarian sentences containing multiple universal quantifier phrases (QPs). Our experimental data on universal QPs not only broaden the descriptive coverage of Hungarian intonation, but also provide new empirical data for unanswered theoretical questions on the syntax-prosody interface. There are two goals in this study, one descriptive and one theoretical. The former is to better understand the prosody of Topic-Comment structure in Hungarian, a language often considered to have a syntactically rigid Topic-Comment structure (É. Kiss 1995). Focus moves to an immediately preverbal position, a designated focus position above TP (Spec,FocP, Brody 1990). Fronted DPs and referential PPs preceding it constitute syntactic topics (Specs of TopPs). Universal QPs, however, can only move to a projection above FocP, but lower than TopP (DistP, Szabolcsi 1997). Thus, we examine whether they prosodically behave as part of Topic or Comment. The second goal is to clarify unanswered theoretical questions regarding the syntax-prosody interface of Focus. Szendrői (2003) claimed that Focus moves to the immediately preverbal position for a prosodic reason, that is, to receive the nuclear stress (NS). Her analysis predicts NS and Focus always coincide at the preverbal position. Universal QPs, however, are unable to occupy the preverbal focus position, but can be the semantic Focus in an appropriate context. The relation between nuclear and focal prominence calls for further investigation. Furthermore, Szendrői assumes that all the pre-focal phrases adjoin to FocP (her extended VP) and become prosodically extrametrical, be it topic or not. It is unclear how Topic-Comment partition is prosodically encoded.

Experiment: All the stimuli have the structure in (1): a scene-setting PP (interpreted as a topic), two universal QPs, followed by a verbal particle in the preverbal position (PP-Q1-Q2-Prt-V). Each sentence is inserted in four different contexts, where the Focus of the sentence is either (a) the entire sentence, i.e. all-new (b) Q1, (c) Q2, or (d) Prt-V.

Results & Discussion: Our preliminary results from 3 speakers are summarized below: (i) In the all-new context (1a), the NS consistently falls on the H*L on Q1, followed by downstepped !H*L accents on Q2 and Prt-V (Fig. 1). This suggests that the default location of the NS is the (highest) QP, if any, rather than the immediately preverbal phrase, as Szendrői claims. (ii) When Focus = Q1 (1b), post-focal phrases (Q2/Prt-V) are either compressed (very small H*L peaks), or deaccented (an L followed by an H%) (Fig. 2). This suggests that an H*L accent is identified as a focal accent only when followed by the post-focal effects, i.e., not all H*L’s are linked to Focus. (iii) When Focus = Q2 or Prt (1c,d), speakers show variation. One speaker shows a clear Topic-Comment partition by shifting the NS from Q1 to the Focus, while assigning various ‘topic contours’—rise (L*H/L*+H%), fall (H*-L%), or plateau (H*)—to pre-focal phrases, be it a PP or a QP (Fig. 3). The other two speakers show a mismatch between nuclear and focal prominence: The NS remains at Q1, while the Focus is marked by a downstepped !H*L accent (Fig. 4), a fact also unexpected in Szendrői’s analysis.

Analysis: We propose that in a default contour, all DistPs and the FocP (extended projections of VP, É. Kiss 2008) are recursively mapped as an intonation phrase (ip), and that the leftmost phrase of each ip bears a default H*L accent, with NS assigned to the head of maximal ip (=2)). We also claim that the inter-speaker variation in (iii) comes from choice between the two principles to obey, which conflict in DistP: Stress-Focus Correspondence Principle (Reinhart 1995), which demands that Focus always bear the NS, and the mapping principle proposed above, which assigns NS to the maximal ip-head (i.e. highest Spec,DistP). Intra-speaker variations, omitted here for space limitation, will also be discussed.
Examples

(1) PP Q1 Q2 Prt-V Post-V
   A vizsgán mindenki mindent megoldott egy óra alatt, ...
   the exam.at everyone everything.acc Prt solved one hour through
   ’At the exam, everyone solved everything within an hour,
   a. All-New Focus (control)
      ..., aztán hozzám. ‘..., and then went home.’
   b. Focus - Q1 mindent ‘everyone’
      ..., nem csak János. ‘... EVERYONE ..., not only John.’
   c. Focus - Q2 mindent ‘everything.acc’
      ..., nem csak a könnyű feladatokat. ‘... EVERYTHING ..., not only the easy tasks.’
   d. Focus - Prt-V megoldott ‘Prt solved’
      ..., nem csak megpróbáltott a megoldással. ‘... SOLVED ..., not just tried to solve.’

(2) [q1] Q1 [q2] Q2 [top Prt-V ...]] ⇒ (q1 Q1 (q2 Q2 (q4 Prt-V ...)))

Figures

Fig. 1. Default contour in an all-new context (1a), by Speaker A: Topic L'H accent on PP, nuclear H*L on Q1, and downstepped H*L's on Q2 and Prt-V.

Fig. 2. Focus on Q1 (1b), by Speaker A:
Nuclear H*L accent on Q1, followed by the post-focal descending.

Fig. 3. Focus on Q2 (1c), by Speaker A:
Topic L'H accents on PP and Q1, followed by the nuclear H*L on Q2 and post-focal compression.

Fig. 4. Focus on Q2 (1c), by Speaker B:
Nuclear H*L accent on Q1 followed by a downstepped (but not post-focal-compressed) H*L on Q2.

References

What information does the intonational lexicon contain? If we presuppose decomposition, accepting the concept of intonational morphemes, then one could perhaps infer that these morphemes are stored with appropriate phonological and semantic information. If so, one could for instance assume that a question-TUNE would have phonological as well as semantic information. What this is so, what does a speaker have to do to prepare to utter a question? In psycholinguistic models of language production, it is often understood that a speaker’s plans to verbalize an utterance involves accessing the conceptual structure, which is followed by accessing the lexicon and encoding the phonological information (Levelt 1989). We could then ask the following: when planning the intonation, does a speaker consider a TUNE to consist of only phonological information, or only semantic information, or both? In this talk, I will examine the interaction of phonological and semantic information for question and focus intonational morphemes, which will include a discussion on the interaction of focus clitics with question tunes, and provide initial experimental evidence on the time course of the activation of the semantic and phonological (intonational) information.
Poster presentations
Factoring out Speaker Variation in Experimental Studies of Prosody: The Case of Association with Focus

M. Breen, M. Wagner, S. Shattuck-Hufnagel, E. Flemming, & E. Gibson
UMass Amherst, McGill University, MIT
mbreen@psych.umass.edu

A limitation on research into prosody has been the difficulty of finding a way to evaluate hypotheses quantitatively. This task is difficult because individual variation in speakers’ productions is often large enough to wash out experimental effects. This paper has two goals: 1) to introduce new statistical methods which account for variability among speakers and identify acoustic features of information structure in an objective way; and 2) to apply these methods to data which investigate how speakers prosodify cases of multiple foci, where the focus operator could associate with one or both of the foci (cf. Krifka, 1992; Rooth, 1996).

In ‘Association with focus’ (Jackendoff, 1972) a sentence’s meaning changes with the prosodic realization of material in the scope of ‘focus-sensitive operators’, e.g. exclusive only. A phrase may contain multiple foci, for example the focus of ‘only’ and a contrastive focus (e.g., condition B in Table 1). Little is known about the prosody of multiple foci, e.g., it is not clear whether listeners can distinguish the associating focus from contrastive focus, given that both are likely to be accented.

We recorded 10 pairs of naïve subjects producing semantically ambiguous target sentences like Grandma only gave a bunny to Maryanne after reading disambiguating contexts (Table 1), with the goal of inducing the listener to select the appropriate picture in Fig. 1. 24 acoustic measures of duration, pitch, and intensity were extracted from 5 target words (e.g. Gramma, only, gave, bunny, Maryanne). Without accounting for speaker and item variation, none of the conditions were discriminated by these measures in pair-wise comparisons. To remove variance due to speakers and items, we computed linear regression models in which speaker (n = 20) and item (n = 20) predicted the 24 acoustic features. From each model, we calculated the predicted value of each acoustic feature per item per speaker. The difference between the predicted and actual values (i.e. the residual measure) reflects acoustic differences due only to experimental manipulation. We submitted the residual measures to a stepwise discriminant function analysis, to independently determine which acoustic measures speakers used to differentiate productions.

Eight acoustic measures (duration, mean pitch, pitch range, and maximum intensity from bunny and Maryanne, respectively) resulted in better-than-chance 6-way classification of the productions according to context; moreover, many conditions were now discriminated in pair-wise comparisons. For example, (B) and (D) were discriminated, although in both conditions bunny and Maryanne are accented and invoke alternatives. The contrastive NPs are longer, have higher pitch, a larger pitch range, and higher intensity than the focused NPs associating with only. Thus contrastive material is more prominent than focused material. Other conditions were also successfully discriminated, including (D) and (E), indicating that previous mention results in a less prominent realization of an NP, even where all NPs are accented (cf. Bard, 2000).

In summary, this paper will present both methodological and empirical results: First, we will describe a way of removing inter-speaker variability in order to reveal hidden systematic patterns in prosody. Second, we will describe how using these methods reveals a set of differentiations among types of foci (see Table 1).
Figures and Tables

Figure 1. Picture display from the production experiment. The Listener selected the picture which best represented the meaning of the target sentence, as produced by the Speaker.

<table>
<thead>
<tr>
<th>CONTEXT</th>
<th>TARGET</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Gramma didn’t give a scarf to Maryanne.</td>
</tr>
<tr>
<td>B</td>
<td>Gramma gave a scarf and a bunny to John.</td>
</tr>
<tr>
<td>C</td>
<td>Gramma didn’t give a bunny to John.</td>
</tr>
<tr>
<td>D</td>
<td>Gramma gave a scarf to both Maryanne and John.</td>
</tr>
<tr>
<td>E</td>
<td>Gramma picked one present and gave it to her favorite grandchild.</td>
</tr>
<tr>
<td>F</td>
<td>Gramma didn’t give a scarf to Maryanne, and she didn’t give either a bunny or a scarf to John.</td>
</tr>
<tr>
<td></td>
<td>Gramma only gave a bunny, to Maryanne.</td>
</tr>
<tr>
<td></td>
<td>Gramma only gave a bunny, to Maryanne.</td>
</tr>
<tr>
<td></td>
<td>Gramma only gave a bunny, to Maryanne.</td>
</tr>
<tr>
<td></td>
<td>Gramma only gave a bunny, to Maryanne.</td>
</tr>
<tr>
<td></td>
<td>Gramma only gave a bunny, to Maryanne.</td>
</tr>
</tbody>
</table>

Table 1: Contexts and focus structures for target sentences used in the experiment. Speakers produced only the target sentences aloud. All Contexts were preceded by a longer Set-up in which, in all but E, all discourse entities (i.e. bunny, scarf, Maryanne, & John) were mentioned. Subscripts are as follows: G = given; F = associates with ‘only’; C = contrastive, but non-associating, focus.

References

Non-local pitch range relationships in read and elicited speech

Alejna Brugos
Boston University
abrugos@bu.edu

The Strict Layer Hypothesis (Selkirk, 1984; Nespor & Vogel, 1986) stipulates that all utterances are parsed into Full Intonational phrases (IPs), which in turn are parsed into one or more intermediate phrases (or equivalent). Theory-based annotation systems, such as ToBI (Beckman et al., 2005), further provide definition of tonal and timing-related properties of intermediate phrases and full IPs. The resulting system leaves little theoretical space for an intonationally defined grouping of prosodic units that are analyzed as Full IPs. Many researchers have observed, however, that tone relationships appear to extend beyond the boundaries of a local prosodic domain, especially in cases where an utterance consists of multiple intermediate phrases, i.e. where the syntactic and/or semantic content of a sequence of phrases is tightly grouped (Ladd 1988, 1992; van den Berg et al. 1992, inter alia).

In experiments with British speakers, Ladd (1988, 1992) showed that in multi-clause utterances, the degree of partial reset (measured via the f0-scaling of clause-initial accent peaks) was sensitive to whether clauses were grouped more tightly, via the coordinating conjunction and, or less tightly, as set off by but (see Figure 1 and Table 1). Truckenbrodt & Féry (2003) and Féry & Truckenbrodt (2005) found similar results for German, further suggesting also that reset values of sentences with the order but/and are comparable to those of sentences without additional branching structure, as with clause-internal downstep, or sentences with clauses with only and-type conjunction.

This study seeks to reproduce results from Ladd, and Féry & Truckenbrodt, and to test the hypothesis that these patterns will hold when additional syntactic and prosodic complexity is introduced. It is hypothesized that speakers will be capable of producing patterns sensitive to structure that extend beyond the confines even of Full IPs, testing the limits of stricter versions of the Prosodic Hierarchy. Pilot data from 4 speakers of American English, who produced sentences of 3 conjoined clauses of comparable syntactic structure and length. Sentences varied by conjunction order (and/but vs. but/and, or and only) and by length (simple or branching subject, simple or branching object) (see Table 2). The introduction of branching subject and object nodes was intended to increase the likelihood of prosodic breaks within the clauses, as well as between clauses. Subjects read sentences aloud and produced sentences comparable to those in the reading task using a picture-based elicitation. The f0 of clause-initial accent-related peaks for 221 tokens was measured.

Comparison of f0 values from clause-initial peaks of the first and third clauses (A and C, see Figure 1) support results from previous studies: the average between-peak difference is smaller for the and/but order than for but/and in all length conditions (Table 3) in both read and picture-elicited productions (Table 4). This pattern held in spite of varying degrees of prosodic complexity within the clauses, as more complex sentences (especially those with both subject and object branching) were more likely to include additional prosodic breaks. Sentences produced via the picture elicitation also contained frequent disfluencies (e.g., within-clause pauses), yet the same reset pattern was observed.

These results support the notion that the phonetic implementation of utterances is sensitive to syntactic complexity and meaning relationships between phrases. The notion that these sentences are produced as a single complex prosodic unit, containing multiple Full IPs, calls for reanalyzing aspects of prosodic hierarchies (Selkirk, 1984; Nespor & Vogel, 1986), such as with recursion of prosodic levels, as suggested by Féry (2008) and Wagner (2005). Such results also suggest a need for expansion of prosodic annotation systems to allow for non-local tone relationships.
Figures and Tables

Figure 1: The structure of 3-clause sentences with varying conjunction orders

a. but/and
b. and/but
c. and only

Table 1: Examples from Ladd (1988)

<table>
<thead>
<tr>
<th>Conjunction/structure</th>
<th>Sample sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>but/and</td>
<td>Ryan has a lot more money, but Warren is a stronger campaigner, and Allen</td>
</tr>
<tr>
<td>A but [B and C]</td>
<td>has more popular policies.</td>
</tr>
<tr>
<td>and/but</td>
<td>Allen is a stronger campaigner, and Ryan has more popular policies, but</td>
</tr>
<tr>
<td>[A and B] but C</td>
<td>Warren has a lot more money.</td>
</tr>
</tbody>
</table>

Table 2: Examples of sentence types from this study:

<table>
<thead>
<tr>
<th>Conjunction</th>
<th>Complexity</th>
<th>Sample sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>but/and</td>
<td>Simple subject, simple object</td>
<td>The bunny got a lemon, but the rhino got an onion and the mummy got a flower.</td>
</tr>
<tr>
<td>A but [B and C]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>And/but</td>
<td>Branching subject, branching object</td>
<td>The baby in the bunny costume won a flower and a bunch of balloons and the baby in the rhino costume won a ribbon and a box of crayons, but the baby in the mummy costume won a lemon and a bowl of berries.</td>
</tr>
<tr>
<td>[A and B] but C</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Average difference in semitones between A1 (f0 of initial accent-related peak of the first clause) and C1 (f0 of initial accent-related peak of the third clause), shown by sentence length and conjunction order, for all 4 subjects.

<table>
<thead>
<tr>
<th></th>
<th>and only</th>
<th>but/and</th>
<th>and/but</th>
<th># of tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple subject/simple object</td>
<td>5.00</td>
<td>4.94</td>
<td>4.65</td>
<td>N=55</td>
</tr>
<tr>
<td>Simple subject/branching object</td>
<td>3.98</td>
<td>4.19</td>
<td>2.71</td>
<td>N=55</td>
</tr>
<tr>
<td>Branching subject/simple object</td>
<td>3.81</td>
<td>3.19</td>
<td>2.56</td>
<td>N=56</td>
</tr>
<tr>
<td>Branching subject/branching object</td>
<td>2.11</td>
<td>3.09</td>
<td>1.41</td>
<td>N=55</td>
</tr>
<tr>
<td>TOTAL</td>
<td>3.72 (N=93)</td>
<td>3.86 (N=64)</td>
<td>2.83 (N=64)</td>
<td>N=221</td>
</tr>
</tbody>
</table>

Table 4: Average difference in semitones between A1 (f0 of initial accent-related peak of the first clause) and C1 (f0 of initial accent-related peak of the third clause), shown by elicitation method, for all 4 subjects.

<table>
<thead>
<tr>
<th></th>
<th>and only</th>
<th>but/and</th>
<th>and/but</th>
<th># of tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read</td>
<td>3.87</td>
<td>3.78</td>
<td>2.85</td>
<td>128</td>
</tr>
<tr>
<td>Elicited</td>
<td>3.38</td>
<td>3.96</td>
<td>2.82</td>
<td>93</td>
</tr>
<tr>
<td>TOTAL</td>
<td>3.72 (N=93)</td>
<td>3.86 (N=64)</td>
<td>2.83 (N=64)</td>
<td>N=221</td>
</tr>
</tbody>
</table>
References


A More precisely, the topic (Fred in (1)) introduces the set of alternative sets introduced by the focus.

A fundamental difference between the two notions is that according to the former, the topic has referential properties, while under the latter, the topic is propositional.

The construction called ‘Clitic Left Dislocation’ (CLLD) – an argument displaced to sentence initial position and resumed by a clitic inside the clause – is a typical construction in Romance languages. Prosodically, the dislocated element (il capo, in 2) is assumed to be set apart in its own phrase (cf. Feldhausen 2008 for Catalan, Frascarelli & Hinterhoelzl 2007 for Italian).

From a pragmatic point of view, the dislocated element represents the topic, namely the referent about which the sentence provides some information (the comment) (Reinhart 1982).

As observed by McNally (1998), a topic is also defined in a different way in the literature, namely as an element that introduces a set of alternative propositions in the semantic computation (Büring 1997, 2003). This ‘contrastive’ notion of topic seems to well fit English and German data. Ain which the topic is assumed to be indicated by a particular accent the so-called see, which is actually a rise-fall-rise tune analyzed as a L+H* accent plus a L-H% edge tone).

It is generally assumed (cf. e.g. Vallduví & Engdahl 1996) that the B-accent in Germanic and the CLLD in Romance express the same discourse function. If that was the case, however, then a uniform definition of topic should account for both constructions. Some attempts have been made in the literature (e.g. Brunetti 2009) to define the function of a CLLD in a way compatible with the definition of B-accent. The idea is that a CLLD too is inherently contrastive (it evokes a set of alternatives). Evidence comes from the exchange in (3), with a partial answer. When an answer is partial, an alternative set must be evoked in order for the answer to be congruent to the question (Büring 1997). In such a context, English requires a B-accent, while in Romance, a CLLD is required (instead of a Clitic Right Dislocation, as in 3C).

Nevertheless, these data do not take into account the prosodic aspects of a CLLD in exhaustive vs partial answers. As a matter of fact, no clear comparison between the two types of answers has ever been made for Romance languages, to the best of our knowledge (as done instead in Germanic languages, cf. Braun 2004, Calhoun 2005).

In this paper we present experimental data in Neapolitan Italian (NI) that show a clear difference in intonation between a CLLD-ed object topic in a congruent answers and in incongruent (precisely, partial) answers to wh-questions. topic expression in a partial answer as seen for preverbal subject topics by &E & a is not necessarily .PARTE FONETICA, CON DESCRIZIONE DELLE CARATTERISTICHE ACUSTICHE DEI DUE ACCENTI....

The finding of a ‘partial’ tune in Romance provides a solution to the problem of defining topic in a uniform way, by supporting a bi-dimensional model of Information Structure (Vallduví & Vilkuna 1998, Steedman 2000). Within this model, Contrast is an independent dimension from the Topic-Comment dimension, and is expressed by the ‘partial’ tune in NI or the B-accent in English. The topic of the topic-comment partition, on the contrary, is expressed in Romance by the sentence initial element (CLLD-ed object or preverbal subject).
References

(1) a. A: What did Fred eat?  B: Fred ate the BEANS.
  b. {{ Fred ate the beans, Fred ate the peas, Fred ate the carrots, etc.}; {Mary ate the beans, Mary ate the peas, Mary ate the carrots, etc.}; {John ate the beans, John ate the peas, etc.}}

(2) Il capo non lo disturberei. ‘(As for) the boss, I would not bother.’

(3) A: Chi porta le bibite per la festa? ‘Who is bringing the drinks for the party?’
  B: La birra(acc), la porta MARIA. ‘MARIA is bringing the beer’
  C: # La porta MARIA, la birra(acc). ‘MARIA is bringing the beer’

Figures

Fig. 1 A Momo gliel’ho dato io (exhaustive answer)  Fig. 2 A Momo gliel’ho dato io (partial answer)

References

Braun, B. (2004), Production and perception of contrastive and non-contrastive themes in German, PhD Dissertation, Universität Saarland.
Encoding focus in French: phrasing, deaccentuation, tonal patterns or all of them?

Aoju Chen and Emilie Destruel
1Max Planck Institute for Psycholinguistics, 2University of Texas
aoju.chen@mpi.nl, emiliedestruel@gmail.com

How intonation is used to encode focus in French is still a matter of debate. One view is that words are accented primarily for structural reasons instead of pragmatic reasons in French (e.g. Garde 1968). A second view is that (narrow) focus is marked by a large rise or fall whereas the pre-focus sequence is realized with a reduced pitch range and the post-focus sequence is realized with a flat contour or deaccented (e.g. Di Cristo & Hirst 1993, De Cristo 1998). Similar observations are made for contrastive focus by Jun and Fougeron (2000), who also suggest that the exact tonal patterns do not matter (2002). There is yet a third view. Skeptical about the relevance of accentuation in French intonation, Fery (2001) argues that French only uses phrasing to encode focus. The co-existence of such diverse views is perplexing and calls for further investigations. In this study, we aim to shed new light on this matter by examining the role of multiple cues in encoding focus of different types and in different positions in a sentence in French. The cues include phonetic cues (i.e. pitch range and word duration), deaccentuation, variations in tonal patterns, and phrasing.

We tested six monolingual speakers of a southern variety of French. We obtained 24 SVO sentences from each speaker in a semi-spontaneous setting by means of a picture matching game. The sentences were uttered as answers to false guesses about the pictures with the correction on the object (Contrastive focus on object or CF-O), answers to WHAT-questions, with focus on the object (NF-O), and answers to WHO-questions, with focus on the subject (NF-S) (see Appendix). Subject and object nouns were all disyllabic words. Each object noun occurred in all focus conditions. Subject nouns were identical in NF-S and NF-O. Each sentence was annotated for tonal patterns and phrasing following Jun and Fougeron (2000, 2002), and acoustically annotated for phonetic analysis in focused words.

In respect of phonetic cues, mixed-effect modelling showed that duration and pitch range appeared to be longer in NF-S than in NF/CF-O but the differences did not reach statistical significance, indicating substantial between-speaker variations in the use of these cues, contra related claims in earlier literature. Regarding post-focus deaccentuation, our binary logistic regression analysis showed that the contour following the verb was significantly more frequently flat in NF-S than in CF/NF-O (P < 0.05), according with Di Cristo & Hirst (1993) and Jun & Fougeron (2000). As regards phrasing, subject nouns almost always formed their own accentual phrases (AP). Interestingly, object nouns formed independent APs significantly more frequently in CF-O (74%) and NF-O (63%) than in NF-S (24%) (p < 0.0001). Finally, regarding variations in tonal patterns, subject nouns were frequently realized with LH* and LLH* across focus conditions. However, in the object nouns, a tonal pattern with a high tonal target (i.e. Hi) occurred significantly more frequently in CF-O (37%) and NF-O (27%) than in SF-O (8.7%) (P < 0.005), contra Jun and Fougeron (2002).

To conclude, our results show that French uses different types of phonological cues in tandem to mark focus. Notably, as in many languages, focus realisation is more pronounced in sentence-final position (phrasing & tonal patterns) than in sentence-initial position (post-focus deaccentuation) in French. Future research is needed to verify the current findings in other varieties of French.
Examples

(1) A sample trial with contrastive focus on the object (CF-O):

Experimenter: Regarde! Un canard! (Look! A duck!)
Est-ce que le canard mange une fraise ? (Is the dark eating a strawberry?)
Participant: (non) Le canard mange un marron.

(2) A sample trial with narrow focus on the object (NF-O):

Experimenter: Regarde! Un lézard! (Look! A lizard!)
Qu'achète le docteur ? (What is the lizard eating?)
Participant: Le lézard mange un marron. (The lizard is eating a chestnut.)

(3) A sample trial with narrow focus on the subject (NF-S):

Experimenter: Regarde! Un marron! (Look! a chestnut.)
Participant: Qui mange le marron ? (Who is eating the chestnut?)
Un lapin achète le marron. (A rabbit is eating the chestnut.)

References

The prosody of information structure in Paraguayan Guaraní

Cynthia G. Clopper and Judith Tonhauser
The Ohio State University
clopper.1@osu.edu, judith@ling.osu.edu

It has long been recognized that information structural meaning can be expressed by prosody (e.g. Paul 1888, Halliday 1967, Chomsky 1972). Recent experimental research has confirmed this relationship, and demonstrated that languages vary in the prosodic phenomena that are used to mark information structural differences. For example, broad and narrow focus can be distinguished by categorical pitch accent differences in Neapolitan Italian (D’Imperio 2002), f0 peak alignment in Spanish (Face 2001), and word duration in English (Eady et al. 1986). In this paper, we explore how information structural differences are realized prosodically in Paraguayan Guaraní (henceforth Guaraní), an under-represented language in the literature on prosody and meaning. Gregores and Suárez (1967: 76), the sole reference grammar of Guaraní, briefly addressed the semantics of Guaraní prosody on the basis of impressionistic transcription and noted that a kind of contrastive focus they described as “emphasis” may be marked by a more extreme pitch contour.

Eighteen native speakers of Guaraní were recorded in San Lorenzo, Paraguay, answering one question for each of nine pictures from Mercer’s *A boy, a dog and a frog* picture book. The target utterances represented three different information structural conditions: subject old/verb contrastive (1a), subject contrastive/verb old (1b), and subject old/verb new (1c), with three target utterances per condition. The analysis was based on the 84 responses that had either subject-verb or subject-verb-PP sentence structure (e.g. (1a) and (1b), respectively) and were composed of a single intonation unit. Responses that did not include both a subject and a verb, had a different sentence structure, an intonation break or pause, or verb-subject word order were excluded. This study is the first acoustic-phonetic exploration of information structure in Guaraní (but see Velázquez-Castillo (1995) and Tonhauser & Colijn (to appear) for discussions of the effects of incorporation and word order on information structure).

The analysis revealed that the stressed syllable of the subject in the subject contrastive condition was significantly longer in duration than the stressed syllable of the subject in the verb contrastive and verb new conditions, and the stressed syllable of the verb was significantly longer in the verb contrastive and verb new conditions than in the subject contrastive condition. These results suggest that contrastive and new focus lengthen the stressed syllable of the focused element in Guaraní. Utterances in all three conditions were realized with an intonation contour consisting of an initial valley on the stressed syllable of the subject, followed by a rise, and then a fall onto the stressed syllable of the verb, as shown in Figure 1. The three information structure conditions did not differ in terms of the alignment of the valley relative to the stressed syllable of the subject, or the duration, f0 range, or slope of the rise from the valley to the peak, suggesting that information structural differences due to contrastive and new focus are not realized by a change in contour in Guaraní. However, the peak of the contour occurred more often on the subject in the subject contrastive condition and more often on the verb in the verb contrastive and verb new conditions. Peak alignment relative to the stressed syllable of the subject was significantly correlated with the duration of the stressed syllable of the subject, suggesting that the effect of peak alignment is due primarily to the longer duration of the subject in the subject contrastive condition.

Taken together, these results suggest that contrastive and new focus are marked in Guaraní by the duration of the stressed syllable of the focused element and, in the case of the subject contrastive condition, that this vowel lengthening has the additional effect of aligning the peak earlier, often on the subject. We did not observe any effect of focus on f0 range, however, in contrast to the observations reported by Gregores and Suárez (1967).
Examples

(1) Stimulus questions (A) and sample answers (B) for the three conditions:
(The experiment instructions introduced the boy as *Fernando* and the frog as *Ju’i*. Glosses used are 3 = third person and QU = question.)

a. Subject old/verb contrastive
Fernando 3-walk or 3-go bike-on Fernando 3-walk
‘A: Is Fernando walking or riding a bike? B: Fernando is walking.’

b. Subject contrastive/verb old
who-QU 3-sit leaf-on Fernando or Frog? Frog 3-sit leaf-on
‘A: Who is sitting on the leaf, Fernando or Frog? B: Frog is sitting on the leaf.’

c. Subject old/verb new
A: Mba’ê-pa o-japo Ju’i? B: Ju’i o-guapy ita-ári.
what-QU 3-do Frog Frog 3-sit stone-on
‘A: What is Frog doing? B: Frog is sitting on a/the stone.’

Figures

Figure 1. Pitch track of the utterance *Fernándo o-guata* ('Fernando is walking') in the subject old/verb contrastive condition, with a valley on the stressed syllable of the subject *Fernándo*, followed by a rise to a peak on the second syllable of the verb *o-guata*, and then a fall onto the final stressed syllable of the verb *o-guata*.

References


Variations on Contrastive Topic Marking - Evidence from Mandarin Chinese
Noah Constant
University of Massachusetts, Amherst
constant@linguist.umass.edu

This article gives evidence for two claims about contrastive topic (CT) across languages. First, as recent work suggests, a unified pragmatic account of CT is possible, covering both “lone CT” (Ward & Hirschberg’s “rise-fall-rise”) as well as CT+F contours, given in (1) and (2) respectively (Wagner 2008, Oshima 2008, Constant 2009). Second, following Steedman (2000), CT meaning can be realized entirely at the boundary of the intonational phrase. Specifically, I show that Mandarin CT constructions have the particle ne in the position of the English boundary tone L-H%. I follow Chao (1968) in factoring out aspectual uses of ne, but show that both “evaluative” sentence final ne (Li 2006) and “thematic” ne (Wu 2006), given in (3) and (4), can convey CT, corresponding to lone CT and CT+F meanings respectively. In contrast, Japanese and Korean CT markers wa and (n)un, are known to occur on the focused phrase, like the English pitch accents L*+H or (L+)H* (Heycock 2007, Lee 2006).

To compare CT realization across languages, it’s important to be precise about what counts as a CT meaning. I show that use conditions on Mandarin ne fit with Büring’s (2003) theory that CT marks a response to a question which is part of a larger discourse strategy, delimited by the CT value of the response. Furthermore, ne displays a characteristic feature of CT in resisting “alternative-dispelling” focuses (Constant 2009). These properties are illustrated in elicited examples (4) and (5). In these examples, the constituent focused by CT is not marked in any special way other than the pitch expansion and accentual lengthening associated with Mandarin focused material generally (Xu 1999, Chen 2002).

CT particles like ne open up new possibilities for exploring CT meaning and distribution. Of 527,794 Mandarin product reviews on amazon.cn, 3–4% contain the particle ne. Corpus examples like (6) reveal that ne is frequently used in questions, which is surprising given the claim in Büring 2003 that CT in questions is impossible. Since English CT is found across a range of speech acts including imperatives like (7), and CT is attested in questions for Czech (Sturgeon 2006), Japanese and Korean (Tomioka 2008), and Mandarin, I argue that English in fact possesses CT questions, but they are not distinguished prosodically from other question types. Since questions are already specified for a boundary tone, the homophony of CT and non-CT questions is evidence for Steedman’s (2000) proposal that English CT meaning resides with the boundary tone.

Although the core pragmatics of CT is shared across language-specific realizations, there are a few minor variations in meaning that are worth mention. First, I observe variation in the ability to use CT on specific types of non-sentential answers to yes-no questions. While CT never marks a resolving answer like (8a), “maybe” answers can be CT marked in English (8b) and Mandarin, but cannot be wa-marked in Japanese. Furthermore, both Japanese and Mandarin resist “rhetorically non-resolving” uses of CT as in (8c).

Another important point of variation concerns the ability to use CT marking on the answer to a strategy-final subquestion. While English lone CT resists issue-resolving uses (Constant 2009), CT+F configurations like the one in (2) are possible as resolving answers. In Mandarin, however, examples like (9) show that even lone CT may be used strategy-finally. This constitutes additional evidence that the collapse of lone CT and CT+F is justified.

The differences in CT distribution across languages suggest the need for a typology of CT meanings. Differences notwithstanding, I have provided evidence that the basic pragmatic effect of CT is the same regardless of phonological realization in terms of prosody or particle. This finding highlights the potential for using discourse particles as a window into intonational meaning.
Examples

(1) (Who’s coming?)
[Persephone] L+H

(2) (Who ate what?)
[Persephone] L+H

(3) (Is Zhangsan going to the conference?)

(4) dåbāfēn de shì qǐng (ne) dōu xiāng sān gè rén yǐkǔài lái zuò.
most all need three person together to do
‘Most of these things need to be done by three people together.’

(5) suǒyǒu de shì qǐng (tne) dōu xiāng sān gè rén yǐkǔài lái zuò.
all all need three person together to do
‘All of these things need to be done by three people together.’

(6) nǐ guān tāmen, shéi guān nǐ ne?
you look after them who look after you?
‘If you look after them, then who’s going to look after you?’

(7) (Come on… ) At least study for [English]...

(8) (Do you think they’ll invite him?)

a. # [I Yes I No I Definitely I ]...

b. [Maybe | Possibly | Probably I ]...
(but I’m not sure)

c. [Of course | Obviously | Duh I ]...
(shouldn’t you know as much?)

(9) nǐ bù-shí qū xīfú (ne), shì qū nà diǎn qián (ne), dūi-bù-duì?
you not-be marry wife ne be marry that little money ne right-not-right
‘You’re not marrying a wife, you’re marrying the money, aren’t you?’

References


Numerous studies have shown the influence of prosody in syntactic parsing, e.g. in the resolution of syntactic ambiguities (Schafer, Carter, Clifton and Frazier, 1996; Snedecker and Trueswell, 2003; Teira and Igoa, 2007). These investigations have shown that prosodic cues can indeed help convey the meaning of ambiguous constructions. In our work we concentrate on the structural ambiguity of restrictive relative clauses (RCs) in sentences such as: Someone shot the servant of the actress who was on the balcony. The syntactic ambiguity lies in the fact that both NPs in the main clause (NP1 the servant and NP2 the actress) are potential antecedents of the subject in the RC. The chosen antecedent reveals the attachment of the RC to either NP1 or NP2 (high vs. low attachment, respectively). Cross-linguistic variability in the attachment preference of ambiguous RCs has been found in previous research (references in Fernández, 2002).

In order to explain such variability, Fodor (1998) formulated a prosodic hypothesis, according to which a universal prosodic processor packages the incoming string of words into chunks at an initial stage of processing, thus leading the parser into the resolution of ambiguities. Fodor proposed the Antigravity Law, a principle which states that the prosodic weight of a constituent (i.e. its length) determines its site of attachment. According to this law, long (heavy) prosodic constituents tend to attach high in the syntactic structure, while short (light) prosodic constituents need to be attached to the previous NP2 in order to form a prosodic constituent, therefore resulting in low attachment. Production experiments by Maynell (1999) and Lovrić, Bradley and Fodor (2000, 2001) report the correlation between sites of prosodic breaks or boundaries and attachment preference. Their data show that a prosodic break before the RC correlates with high attachment (NP1 NP2 / RC), while the presence of a break between the two nouns of the complex NP indicates low attachment (NP1 / NP2 RC).

Previous behavioral studies report a high attachment preference in Spanish (Cuetos and Mitchell, 1988; Carreiras and Clifton, 1993, 1999). In our investigation, we analyze the prosodic patterns of these structures in Spanish through a production experiment with three Spanish monolingual speakers from the Basque Country and six Spanish-Basque bilingual speakers (three L1 Spanish and three L1 Basque). In order to test the Antigravity Law, we manipulated the length of the RCs and created four different types of RCs, by length: 3-4 syll., 6-7 syll., 9-11 syll., and 14-15 syll. There were 12 sentences per each degree of length, for a total of 1296 utterances. The speakers had to read the sentences as naturally as possible, three times. We analysed the prosodic contours of these sentences, looking for cues of prosodic boundaries between NP1, NP2 and RC. The prosodic breaks after NP1 or NP2 consisted of continuation rises, sometimes followed by pause and/or final lengthening.

The results of the 1153 sentences analyzed support the Antigravity Law proposed by Fodor. Our data reveal a correlation between length of the RC and site of the prosodic break, showing a higher percentage of prosodic boundaries after NP2 – hence high attachment- in longer RCs, than in short RCs. That is, longer RCs would seem to constitute prosodic phrases of their own. This independence would be signaled by prosodic boundary cues to their left, separating NP2 and RC. In addition to that, significant differences between the linguistic groups can be found. Bilingual speakers in both groups produce more and clearer prosodic boundaries than monolinguals, a difference that deserves further discussion.
References


Adults with autism commonly accent speech elements at the beginning of a sentence, whereas typical
speakers show more variation (e.g., Shriberg et al., 2001). Those diagnosed with High-Functioning Autism
(HFA) also show more intonational abnormalities than those diagnosed with Asperger’s Syndrome (AS;
Fine et al., 1991), a form of autism that is not associated with early language delay (American Psychiatric
Association, 1994). The current study investigates (1) how speakers with Autism compare to typical
speakers in the use of intonation (phonologically and phonetically) to encode topic and focus in English
declaratives; and (2) whether adults with HFA differ from adults with AS in this respect. Topic is
operationalised as the referent that a WH-question is about; focus as the information required by the
WH-word (Lambrecht, 1994).

We tested 3 AS speakers, 3 HFA speakers, and 5 typical adults (controls). All participants had normal
hearing thresholds and were monolingual speakers of Canadian English. We obtained 22 SVO
declaratives in a semi-spontaneous setting by means of a picture matching game from each participant
(see Appendix). Half of the sentences were answers to WHO-questions, with initial focus and final topic.
The other half of the sentences were answers to WHAT-questions, with initial topic and final focus. The
intonation of each sentence was transcribed following the principles in ToDI (Gussenhoven, 2005) and
acoustically annotated for phonetic analyses.

We found that controls accented focus in all cases. Multinomial logistic regression modeling showed
that the most favoured accent type was H*L sentence-initially and !H*L sentence-finally. But controls
realised topic differently depending on its position in the sentence. Sentence-initial topic was always
accented and most frequently with H*L, like focus. Sentence-final topic was predominantly realised with
no accent. In line with Horn (1991) and Chen (2007), we argue that accentuation is used for rhythmic
motivation in sentence-initial topic. However, sentence-initially focus and topic are distinguished via
phonetic parameters, that is, focus had a larger F0 excursion and a longer word duration than topic
across accent types and within nouns spoken with H*L.

Unlike the controls, the AS speakers’ accent placement and choice of accent type appear to be
completely structurally determined. Independent of information structure, sentence-initial nouns were
always accented (with H*L and H*), whereas sentence-final nouns were most frequently spoken with
!H*L. Interestingly, mixed-effect models showed that focus had a longer word duration than topic across
accent types as well as within nouns spoken with H*L or !H*L. The AS speakers thus relied solely on
duration to encode topic and focus. In contrast, the HFA speakers operated on a strict mapping of accent
placement to information structure. Largely, they realised focus with H*L and topic with no accent,
independent of sentence-position. There were a couple of exceptions to this pattern: the use of H*
(spreader L), and H*L (speaker M) in sentence-initial topic and the use of !H*L (speakers L and M) in
sentence-final topic.

To conclude, speakers with autism can use intonation systematically to encode topic and focus.
However, they differ from the controls in their choice of intonational cues. The controls use both
phonological and phonetic cues. The AS speakers rely on phonetic cues (i.e. duration); the HFA speakers
rely on phonological cues (i.e. accent placement), although individual variations exist in the extent to
which accent placement is used. The HFA speakers’ strategy may be the most effective one in
distinguishing topic from focus but it sets the HFA speakers apart from the controls for not taking
rhythmic motivation into account when executing choices of intonation patterns.
Appendix

1. A sample trial with initial topic and final focus:

Experimenter: Look! A rabbit. [shown picture of a rabbit holding a brush] It looks like the rabbit is painting something. What is the rabbit painting? [shape disappears to reveal a picture of a ball being painted black]

Participant: The rabbit is painting a ball.

2. A sample trial with initial focus and final topic:

Experimenter: Look! A bed. [shown picture of a bed with blue paint on it] It looks like someone is painting the bed. Who is painting the bed? [shape disappears to reveal a picture of a rabbit holding a brush next to a paint can]

Participant: The rabbit is painting a bed.

References

Chen, A. (2007). Intonational realising of topic and focus by Dutch-acquiring 4- to 5-year-olds. Presentation at the International Conference for Phonetic Sciences, Saarbrücken, Germany.
I present novel data in Turkish that describes the prosodic properties of different types of topic. It has been shown previously that a H boundary tone optionally demarcates topics in Turkish (Kamali, to appear). Here, I examine the behavior of contrastive and merely anaphoric Topics in Turkish with respect to this boundary tone and two other factors: what happens in the accentless postverbal area, and how topics are realized in questions.

The H boundary tone in question is shown to be obligatory in the case of contrastive topics, but not in the case of merely anaphoric topics. This H tone is easily spotted (right panel in Figure 1, cf. 2B). The next phrase starts off interpolating this H tone to baseline. Also, the last vowel in the topic is much longer in the contrastive topic (right panel) as opposed to the anaphoric topic (left panel in Figure 1, cf. 1B).

Topics in questions follow the same pattern. An anaphoric topic is not marked by a boundary tone (Lefthand panel, Figure 2, cf. 3A), while a contrastive topic has a H tone associated with its left edge (Righthand panel, Figure 2, cf. 2A).

A contrastive topic in a question is perceived increasingly like a focus. This is reflected in the durations of the topic phrases Ali and Emre in the righthand panels of Figures 1 and 2, respectively. I argue that the H target in the wh-word in (2A) causes a merger of the H topic boundary and H*L pitch accent on ‘what’, and therefore the contrastive topic in questions have a focus-like quality.

Lastly, contrastive topics cannot be in the postverbal domain, while anaphoric topics can. I argue that this is due to a prosodic constraint in Turkish that causes the postverbal domain to be tonally flat. Since contrastive topics must be accompanied by a H tone, they cannot be realized in this toneless domain. The acceptability contrast between (2B) and (2'B) shows this point.

This preliminary study reveals interesting parallels between Korean and Japanese topic prosody and that of Turkish (see Lee 2007 and Nakanishi 2007 for analyses of the respective languages). Turkish and Korean both utilize a H tone to mark contrastive topics but not anaphoric topics. Japanese, on the other hand, utilizes phrasing. I argue that on the assumption that these H tones are phrase boundaries in Turkish and Korean, this cross-Altaic topic typology follows smoothly.

Examples

(1)  A: Ali ne yiyor?  
     Ali what eating
  B: Ali elma yiyor.  
     ‘Ali is eating an apple.’

(2)  A: Peki Emre ne yiyor?  
     Then Emre what eating
     (Emre/Ali: Contrastive topic)
  B: (Emre’yi bilmiyorum ama) Ali elma yiyor,
     (I don’t know about Emre but) Ali apple eating
     ‘(I don’t know about Emre but) Ali is eating an apple.’

(2’) A: Peki Emre ne yiyor?  
     Then, what is Emre eating?
  B: #(Emre’yi bilmiyorum ama) elma yiyor Ali,
     (I don’t know about Emre but) apple eating Ali
     #(I don’t know about Emre but) eating an apple is what Ali is doing,’
Figures

![Figure 1: Topics in Declaratives](#)

![Figure 2: Topics in Questions](#)

References


This presentation is intended to show how accentuation in German is related to argument structure, especially in examples with predicate composition.

In neutral German verb-final sentences, all arguments are accented whereas the verb lacks an accent (cf. (1)).

This fact can be explained by using Phrase Structure Grammar: One argument can be realized as a complement to V. Further arguments as well as modifiers are adjuncts to VP. If we assume that in head-complement structures only the complement receives an accent whereas in adjunct structures both constituents are accented, we get the correct accentuation (cf. fig. 1 below). But there are some cases in which arguments remain unaccented (cf. (2)).

The arguments in (1) stand in no direct semantic relation to each other. The children are neither in the garden during the event nor coming into the garden through the event. In (2a-c) the DP argument stands in a BECOME relation to the following argument. PP, AP, and PTKL are secondary predicates which add a result state to the event and change the situational type from activity to accomplishment. They are realized as complements to V. But instead of building a maximal phrase with V, they merge with V to a composed predicate which can have the DP as further complement (cf. fig. 2 below). The DP is a semantic argument of the secondary predicate (cf. (3)) but is syntactically realized under a projection of V. Arguments of V can be suppressed by predicate composition.

The accentuation can be derived from the resulting syntactic structure in fig. 2 below. In a top-down operating accent assignment process, accent features are given to the DP das Kind and to the VP das Glas leer_trank because they are adjuncts. In a next step, the accent feature of the VP is given to the DP das Glas because it is the complement of V+. The AP as a complement of V cannot get an accent feature because it is dominated by a higher-level head, which blocks the assignment.

In my presentation, I would like to show the conditions for predicate composition and the consequences for their syntactic and prosodic realization in contrast to constructions in which predicate composition is not possible (e.g. with depictive adjectives or non directional PPs).

Examples

(1) weil am MORgen KINder in den GARten schauten
because in-the morning children into the garden looked

(2) a. weil am MORgen KINder in den Garten rannten
because in-the morning children into the garden ran
b. weil das KIND das GLAS leer_trank
because the child the glass empty drank
c. weil das KIND das GLAS um warf
because the child the glass PTKL threw (= knocked over)

(3) a. #weil das Kind das Glas trinkt
because the child the glass drink (V)
b. "weil das Glas leer ist / wird
because the glass empty is / becomes
References:


Three types of prosodic focus in Brazilian Portuguese: form and meaning
João Antônio de Moraes
Universidade Federal do Rio de Janeiro / CNPq
jamoraes2@br.inter.net, chermanny2@gmail.com

It is a well known fact that in intonational languages melodic contours strongly correlate with information structure, especially with the manifestation of focus. This correlation can be achieved in terms of distribution of pitch accents along the sentence, and eventually by the choice of the pitch accent itself, distinguishing different types of focus. Beside the type of focus, several other conditions can also affect the melodic expression of focus, including the extension (number of syllables) of the focused constituent, its accentual pattern, syntactic complexity, word order, and the position of the focused element in the utterance (final vs. non-final).

In this paper I show that in Brazilian Portuguese (BP) there are three different nuclear contours for final position focus, according to different (semantic) types of focus:

(i) informational focus:
A: Que carro Francisco vendeu ontem? (What car did Francisco sell yesterday?)
B: Francisco vendeu \textbf{uma Mercedes} (Francisco sold a Mercedes)
\begin{tabular}{|l|}
\hline
H & L* & L% \\
\hline
\end{tabular}

(ii) exclusive corrective focus:
A: Francisco vendeu um Porsche ontem. (Francisco sold a Porsche yesterday)
B: Francisco vendeu \textbf{uma Mercedes} \textbf{e} (Francisco sold a Mercedes)
\begin{tabular}{|l|}
\hline
¡H & L* & L% \\
\hline
\end{tabular}
(Implies: You are wrong, it is not a Porsche that Francisco sold yesterday)

(iii) non-exclusive (attenuated) corrective focus.
A: Francisco vendeu um Porsche ontem. (Francisco sold a Porsche yesterday)
B: Francisco vendeu \textbf{uma Mercedes} (Francisco sold a Mercedes)
\begin{tabular}{|l|}
\hline
¡H* & L* & L% \\
\hline
\end{tabular}
(Implies: I am sure he sold a Mercedes, I don’t know if he also sold a Porsche, it could be...)

Utterances conveying these 3 types of focus were recorded in a Laboratory set, preceded by the contexts that could elicitate them, and were submitted to an identification test with 15 subjects, who were asked to tell, in a forced choice task, which kind of context had elicitated them, and in the case of exclusive and non-exclusive focuses, their specific meaning. Also 16 resynthesized speech stimuli with modified F0 contours were created and submitted to the evaluation of the same subjects, in order to verify what kind of melodic modification was relevant to the meaning perceived, and in what specific points they occurred. Results show that the distinct focus meanings were in most cases recognized by the listeners and that the phonological analysis proposed was supported by the answers of the identification test with the modified F0 stimuli.
Sentence *Francisco vendeu uma Mercedes*, presenting informational focus on *Mercedes* in panel A, exclusive corrective focus in panel B and non-exclusive corrective focus in panel C.

References


In this paper, we present original data on Beaver (Athabaskan) gathered on six fieldtrips. Beaver is an endangered Athabaskan language with lexical tone spoken by some 150 speakers in British Columbia and Alberta, Canada. Here, we report on two phenomena related to the interaction of prosody and information structure: First, in narrative data, we found a falling initial phrase accent –HL marking contrast. Second, in a study of complex noun phrases, we found that new/contrastive information was produced with a significantly greater pitch range than given information.

It is generally assumed that tone languages are more restricted in their intonational inventory than languages without lexical tone. Especially for focus marking, there are reports on mechanisms such as pitch range manipulation (Mandarin, e.g. Xu 1999), increased duration (Taiwanese, Pan 2007), and phrasing (Bantu, e.g. Hyman 1999) employed in tone languages, but none on intonational tones. Intonational tones marking information structure have only been found in the pitch accent languages of Europe (e.g. Central Franconian, Gussenhoven et al. 2004). For Navajo, an Athabaskan tone language remotely related to Beaver, McDonough 2003 states that information structure is not marked by prosodic means. In this context, the presence of a falling phrase accent –HL marking contrast in Beaver presents an interesting case. –HL can be accompanied by particles such as laa, a focus marker, or zǫ meaning “only”, or it can occur without any additional particles. The phrase containing –HL is produced with an expanded pitch range while pitch range is compressed in the following phrase.

In a controlled study of complex noun phrases (“Animal Task” in QUIS, Skopeteas et al. 2006), conducted with five speakers, we found that new/contrastive information was produced with an increased pitch range for four speakers, while one speaker raised both baseline and top line, using a higher pitch range. In fig. 1, the handlabelled minima and maxima per word have been plotted for each speaker, showing that the maxima are significantly higher in new than in given words for all speakers, while the minima are not effected in the same way.

These findings from Beaver complement current knowledge about intonation in tone languages: On the one hand, providing more evidence for the widespread mechanism of pitch range expansion, while on the other hand, reporting on a phrase accent –HL marking contrast, which is cross-linguistically rare.
Fig. 1: Maxima and minima per word, comparing new/contrastive (white boxes) and given (grey boxes) information for each speaker.

References

Like other Romance languages, Spanish has traditionally been considered to resort mainly to variations in constituent order (Ladd, 1996) to express focal structure. Thus, it may be classified among the ‘non-plastic’ languages (Vallduví and Engdahl, 1996). While some scholars do not think it is possible to shift the position of the nuclear pitch accent to reflect focus (Sosa, 1991), others maintain that such a change is possible only in cases of contrast or emphasis (Zubizarreta, 1998). Among intonation specialists who have worked on the variation in the position of the nuclear accent to reflect focus, Toledo (1989) and Face (2000, 2001, 2002) have explored different phonetic and phonological aspects of focus marking in Buenos Aires and Castilian Spanish respectively, both with a contrastive and a non-contrastive meaning.

This paper explores the prosodic expression of information structure through deaccenting in the Buenos Aires dialect as a phenomenon at the interface between phonology and pragmatics. The corpus consists of spontaneous speech in TV and radio interviews. Spontaneous speech differs from lab speech in several important respects (Face, 2003). Deaccenting is found more frequently in impromptu speech than in lab speech. Despite disfluencies, hesitations, voice overlap, etc, spontaneous speech constitutes a faithful source of what speakers do when they communicate with each other.

The phonological analysis follows the guidelines of the autosegmental-metrical approach to intonation (Pierrehumbert, 1980; Ladd, 1996) applied to Spanish (Sosa and Hualde in Prieto, 2003). It pays particular attention to the significance of the last pitch accent in the intonational unit (Ladd: 204, 1996). The pragmatic analysis is carried out within the framework of Relevance Theory (Sperber and Wilson, 1995, 2004). Relevance is a property of cognitive and linguistic inputs. In communicating, speakers balance considerations of effect and effort: the more cognitive effects an utterance achieves, the more relevant it is; the more processing effort it demands, the less relevant it will seem. Like all other linguistic inputs, prosody contributes to the relevance of utterances by guiding the hearer towards the desired cognitive effects. Any increase in processing effort to achieve those effects is justified only if it results in a global increase of the relevance of the utterance. (Wilson and Wharton, 2006).

As Spanish favours rearrangement in constituent order to express focus by placing focal constituents at the end of the sentence, the nuclear accent usually occurs at the end of the intonational unit. This is the least demanding position from the point of view of processing effort. When the final focal accent is shifted to earlier positions and the final constituent(s) is/are deaccented, the resulting extra processing effort is justified because it achieves cognitive effects which increase the global relevance of the utterance by avoiding a more costly linguistic formulation of the same idea, for example by syntactic rearrangement.

The corpus collected shows that Buenos Aires Spanish allows deaccenting for pragmatic purposes. In fact, it seems to follow two strategies: the post-focal stretch may be completely deaccented (figure 1 below), or post-focal accents may be retained (figure 2 below). A similar observation has been made by Ortiz Lira (1995, 2000) for the Santiago variety of Chilean Spanish. Following García-Lecumberri (1995), these two can be considered to be manifestations of one and the same phonological choice: deaccenting given information in post-nuclear material. Post-focal accents may result from rhythmical considerations, as Spanish seems to prefer to retain word stress in connected speech.
Figures

Figure 1: an example of post-nuclear deaccenting
Figure 2: an example of a post-nuclear pitch accent (marked with an arrow)

References

In English, the reference of free, personal pronouns may depend on the prosodic context in which they occur. In (1), for example, the reference of *him* appears to depend on whether the pronoun receives a nuclear pitch accent (Akmajian & Jackendoff 1970).

In one type of account, the presence of an accent on a pronoun simply “switches” its reference relative to that of some unaccented counterpart (Solan 1983, Kameyama 1999). On this view, accents have an interpretation that is specific to their occurrence on pronouns and independent of information structure (henceforth *IS*). On another view, contrasts like (1) are predictable from the interpretation of accent patterns *within* a model of IS (Venditti et al. 2001, deHoop 2004). This talk reports on a perception study that compares these two classes of theories in cases where their predictions conflict. The findings show that the relationship between prosody and pronominal reference crucially depends on IS, and cannot be accounted for by the notion of a reference switch.

The materials included 32 two-sentence contexts like that in (2), which were presented both auditorily and visually to 32 adult speakers of North American English.

In one condition, the second sentence was produced with no accent on the pronoun as in (2iia). This condition established that the preferred reference of the unaccented pronoun is for the subject of the *matrix* clause of the preceding sentence (i.e., *Alex* in 2i; 66%). According to a “switching” account, therefore, the preferred referent of the accented pronoun (i.e., 2iib) should be the subject of the *embedded* clause (i.e., *Roger*). By contrast, Schwarzschild’s (1999) model of information structure predicts, as I will show, that the accented pronoun in (2iib) cannot refer to the embedded subject of (2i).

In a forced-choice task, subjects chose between two paraphrases of the target sentence (i.e., 2iia/2iib). Forty-four fillers served as distractors and as controls on subjects’ attentiveness. Contra the switching account, accented pronouns actually showed an increased preference for the matrix subject of the preceding sentence (78%, p<0.01). In a related set of conditions, the context (2i) was modified so that, according to Schwarzschild’s model of IS, the accented pronoun was instead consistent only with reference to the *embedded* subject. In this modified context, the accented pronoun showed a strong dispreference for reference to the matrix subject (21%, p<0.001), while the preference in the unaccented case did not differ significantly differ from that in the original context (72%).

Together, these results show that pronominal reference is sensitive to the interpretation of prosody vis-à-vis information structure, even when this contradicts the predictions of a reference switch effect. I show that the analysis of the findings in terms of specific hearer inferences from IS not only extends to classic examples like (1), but is also the more parsimonious account. In addition, I show how this tendency for free pronouns to reflect the structure of local discourse (Kehler 2002) may be used as a discrete detection measure for exploring further aspects of the mapping between prosody and meaning.

**Examples**

(1) John hit Bill, and then George hit *him/HIM*. (*him*=Bill, *HIM*=John)

(2) i. At the hotel, Alex reminded Roger to ask for the executive suite.
   
   iia. Later that night, he made a request.
   
   \[ L+H^* \quad L-H\% \quad H^* \quad L-L\%
   
   iib. Later that night, he made a request.
   
   \[ L+H^* \quad L-H\% \quad H^* \quad L-L\% \]


Figure 1. Proportion of matrix subject responses by prosodic pattern (x-axis) and by context (solid vs. dashed line). Levels of context: A model of information structure predicts that the accented pronoun may refer (i) only to the subject of the preceding matrix clause (dashed line), or (ii) only to the subject of the preceding embedded clause (solid line).

References

L2 learners encounter many difficulties in the acquisition of suprasegmental features, mainly because of L1 influence. Mennen (2006), comparing different studies about intonation in L2 productions, identified a set of characteristics influenced by L1, which she divided into two groups: phonological errors (use of a category instead of another one) and phonetic errors (different realization of the same category).

The work presented here is part of a larger research which aims at exploring the realization of suprasegmental features of German produced by Italian speakers from Lecce (South Italy) with different L2 competence levels (low, middle and high) and from both an acoustic and a kinematic point of view. A first study has already been performed for Wh- and Y/N questions, considering only acoustic semispontaneous materials (Stella, in print): it has been pointed out that the implementation of pitch accents, in terms of alignment of F0 peaks with target syllables, changes gradually in some way through competence levels.

The purpose of the present study is to collect a larger amount of data in order to analyse implementation in a quantitative and statistical way. In this phase, only acoustic data of various focus conditions (broad, narrow and contrastive) produced by six speakers were recorded both in Italian and German. A description of the contours exploited for each focus condition is given, following the Autosegmental-Metrical theory (see Ladd, 1996 for a general review) and using two modified versions of ToBI (Beckman and Ayers, 1997): for Italian, a version of the original ToBI system adapted for the Italian variety of Lecce is used, which has been proposed for the description of the question contours and some focalisation processes (Stella and Gili Fivela, in print); for German productions, GToBI (Grice et alii, 2005) is used, as realisations of Standard German intonation are expected from speakers with higher competence (otherwise variations of GToBI are used in case of non-conformity with Standard German). German productions are also put in relation with L2 competence level and with the phonetic/phonological system of L1 intonation.

For this aim, the realisation of pitch accents on the syllable /ma/ and /man/ is analysed. Indeed, when these syllables are stressed, they are implemented phonetically in the same way for both Italian and German (cfr. Di Meola, 2007) and therefore they are directly comparable (this is important as the corpus will also be used for kinematic acquisitions). For each language, 18 words with the target syllables in stressed position was chosen: they were composed by a variable number of syllables (from 2 to 5 for Italian and from 1 to 5 for German) and had different stress structure (stress on antepenultimate, penultimate and ultimate syllable). Then, they were used to create two corpora of sentences (one in Italian and one in German) and for each sentence a preceding question created the context for focus interpretation. Both corpora were read at least five times by each subject.

As it has been pointed out in Stella (in print) and observed from a first analysis of the data acquired for the present work, an implementation of accents comparable to that of Standard German is difficult also for speakers with higher competence. It could be ascribed to alignment of F0 peaks for the realisation of the same phonological category: this feature is indeed very difficult to learn, as Mennen (2004) pointed out with an analysis of alignment in Dutch speakers of Greek with a native-like competence. As Ladd (1996) suggest, there are many differences in alignment, which can emerge from a comparative perspective. This study aims at shedding light on the difficulties in implementation of accents for L2 speakers beginning from an analysis of alignment, because it seems to be one of the first sources of phonological and phonetic errors also for L2 speakers with higher competence.
References


In traditional phonemic analysis, phonologists rely on speaker intuitions regarding meaning differences to demonstrate phonological contrast in sound patterns. The discreteness of these differences, typically taken for granted, serves as the basis for inferences regarding the discreteness of phonological distinctions. In intonational phonology, however, analogous native-speaker intuitions are often substantially less sharp. Intonational meaning is notoriously slippery, rarely truth-conditional, and frequently confounded by informants with aspects of affect and paralanguage. Native speakers and linguists alike often stumble over seemingly innocent questions, such as whether two contours (e.g., L+H* L-H% and L*+H L-H%) “mean the same thing”. Because of this, the traditional King’s Highway to phonemic contrast, the minimal pair test, is often of little use in the study of intonation, leading researchers [e.g., 5, 9] to seek demonstrations of phonological contrast by other means, such as the identification of sharp boundaries between intonational categories in perception or production. Pierrehumbert [8: 59-63] explicitly advocates discreteness in realization as evidence for phonological contrast, and from there infers distinctions in meaning, rather than the other way around.

Locating categories in perception or production naturally requires a variable expressing the relevant boundaries. In AM intonation models, this usually involves scaling or alignment of F0 turning points (maxima, minima, “elbows”, hereafter TPs) [3,8]. While few would posit exceptionless correspondence between phonological tones and TPs [6: 103-107], most arguments for contrast in intonation today presuppose some level of equivalence. This assumption yields tremendous successes as a laboratory tool [1, et seq.], but nonetheless encounters difficulties as a model of “what matters” in performance. If TP alignment provides a primary cue to tonal representation in phonology, then it should be robustly recoverable in both perception and acoustic analysis. An increasing body of evidence, however, suggests that this is not case; instead, apparently crucial TPs are either difficult to locate, as in ambiguous ‘elbows’ in the F0 contour [2], or absent, as in voiceless regions. Aspects of global contour shape beyond TP alignment are also known to play a role in categorization [4, 7]. This paper presents a new model of tonal alignment that distinguishes contrasting categories in English more robustly than TP-based models, while capturing the influence of contour shape on perception as well. At the heart of this model is the notion of Tonal Center of Gravity (TCoG), a distillation of contour shape returning the location where the “central bulk” of a pitch excursion lies. TCoG is calculated as an average of timepoints in an interval weighted by their F0 values. (See figure 1.) Points with higher f0 are thus more influential then those with lower f0, “pulling” TCoG in their direction.

To investigate the robustness of TCoG as a marker of tonal alignment contrasts, we elicited tokens of L+H* L-H% and L*+H L-H%, contours from 6 native English speakers. Productions were labeled for TPs and TCoG, and automatically categorized by logistic regression analyses representing all models. To explore robustness, we then subjected both TP labels and interval edges for TCoG to random perturbation drawn from a Gaussian source, Classification was thereupon repeated. This procedure was repeated 30 times at 3 levels of added label “noise”. For noiseless data, TCoG categorized better than either the “Low” or the “High” TPs alone, and approximately as well as a model combining these two. With perturbed labels, TCoG significantly outperformed all other models. TCoG is thus less vulnerable to ambiguities in the F0 contour, and we suggest also a more reliable perceptual cue to contrasts between contours similar in shape, but differing in meaning.
**Figures and Tables**

Figure 1: Tonal Center of Gravity (TCoG) is an $f_0$ weighted average of time points, representing the center of the “bulk” or “mass” of $f_0$ over an interval.

$$TCoG = \frac{\sum (\text{time} \times f_0)}{\sum f_0}$$

Table 1: Correct detection of L+H* L.H% vs. L*+H L.H% when the rise and peak are perturbed by Gaussian noise based on multiples of the standard deviation for each speaker’s rise/peak location. Figures represent averages over 30 trials for each noisy condition.

<table>
<thead>
<tr>
<th>Level of noise</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rise (L)</td>
<td>79</td>
<td>70</td>
<td>65</td>
<td>60</td>
</tr>
<tr>
<td>Peak (H)</td>
<td>83</td>
<td>72</td>
<td>66</td>
<td>57</td>
</tr>
<tr>
<td>Rise + Peak (LH)</td>
<td>90</td>
<td>78</td>
<td>70</td>
<td>61</td>
</tr>
<tr>
<td>TCoG</td>
<td>90</td>
<td>85</td>
<td>80</td>
<td>72</td>
</tr>
</tbody>
</table>

**References**

Friday September 18
Oral (Invited and Selected) Presentations
My talk will examine the pragmatic meanings of prosody in context. I will address both intonational phonology and phonetic implementation, and their roles in reflecting and constructing social context.

I will first provide an overview of research into prosody in relation to a variety of social contexts. This will include recent work on regional variation, gender, age and sexual orientation. I will then move on to discuss the interpersonal meanings generated by prosody in social context. How people speak in different situations is intimately bound to participant relationships, and the communicative behaviour that is licensed by these relationships. This is the focus of some of my own recent work (Wichmann 2004, 2005) which has shown, for example, that the prosody of requests varies according to the power relationships between participants.

Finally, I will address the meaning of prosody in social context in relation to the broader phenomenon of social convergence in non-verbal signals. It has been observed that humans tend to display convergent behaviour over time. In terms of prosodic behaviour, this means that, in certain situations, speakers have a tendency to accommodate their pitch to that of their interlocutors, particularly across turns (e.g. Brazil, 1985). The significance of sequential information – the meaning generated by the prosodic relationship of one turn to another rather than by anything inherent in an individual utterance – is beginning to be addressed within the framework of interactional sociolinguistics, e.g. Walker 2004, Culpeper et al 2003, and most recently by Roth and Tobin (to appear). I hope to show that both convergence and divergence are powerful resources in the constructing of relationships in situations of both harmony and conflict.

References

Wichmann A. 2005 ‘Please’ – from courtesy to appeal: the role of intonation in the expression of attitudinal meaning English Language and Linguistics 9 (2) 229-253
Understanding how prosodic patterns of pitch and duration map to meaning is notoriously difficult. Since meaningful distinctions are mediated by phonological representations, accurate phonological theories and models of prosody are essential to advancing knowledge about how prosodic variation conveys meaning in language. In the first part of the talk, I highlight research aimed at studying phonological categories in intonation using a variety of methodological techniques. A central question in this work is the validity of various assumptions made within the well-known autosegmental-metrical (AM) approach to phonology. Of particular interest are the phonological pitch contrasts assumed to underlie meaningful distinctions in AM theory; many of these proposed contrasts are assumed to be distinguished by the alignment of the fundamental frequency (F0) contour relative to the metrical structure of an utterance. While AM theory has been applied to many languages, surprisingly few studies have specifically investigated the validity of the basic, proposed pitch accent contrasts within this theory. I will describe the results of multiple perception and production experiments testing the validity of these pitch accent contrasts using resynthesized speech in which the alignment of F0 extrema was manipulated relative to the metrical structure of an utterance. While these studies generally validate the proposed contrasts, they also suggest the existence of an additional phonological category not assumed to be part of standard, contemporary AM theory. Implications of this category for AM theory and for understanding the mapping of intonation to meaning will be discussed. In the second part of the talk, I focus on results of recent studies which demonstrate unexpected ways in which prosodic patterns of pitch and duration map to meaning. Focusing on English, I will describe several experiments in which the duration and/or F0 of a frame utterance are shown to influence perception of what words are spoken. Since a basic assumption in most prosodic theories is that duration and F0 should not influence understanding of the words that are spoken in languages lacking lexical tone and true phonological gemination, the results provide a puzzling demonstration of how prosodic variation can influence meaning. The implications for phonological theory will be discussed, with an emphasis on how integrating research approaches and findings across disciplines is likely to yield insight into the relationship between prosody and meaning.
It is widely known that one of the functions of prosody is to signal the information structure of an utterance, i.e. its partitioning into Topic, Focus or Given (Background) material. The phonology of intonation can also (among other linguistic means such as word order and morphological marking) convey structural differences such as focus scope and, in some languages, even focus type (cf. Face and D’Imperio 2005 for the difference between contrastive and broad focus in Italian and Spanish). Recent work on German and English intonation has recently questioned the impact of information structure in terms of the topic/comment partition on prosodic patterns. Topic information can be contrastive (Büring, 2003) in that it introduces a set and gives rise to an implicature that if a property holds of the topic, a different property holds of other members of the set that contains the topic (cf. Vallduví and Vilkauna 1998). Moreover, a Contrastive Topic (CT) has been claimed to be cross-linguistically marked by a prominent intonation pattern, i.e. a specific accentual configuration. In German, contrastive topic accents appear to have later and higher peaks (Braun, 2006). In Italian, Frascarelli and colleagues (Frascarelli, 2000; Frascarelli and Hinterholz, 2006) have proposed that all Topics are set aside in their own specific prosodic phrase. We will show here that Topic/Focus partitioning in non-contrastive utterances does not need to be signalled through a phrase break. Moreover, I shall argue that a local prosodic break is not the only means that Italian can exploit in order to signal information structural properties, since the register level of the post-topic domain can be manipulated so as to enhance the contrastive nature of the Topic itself (D’Imperio et al. 2009, D’Imperio, to appear).

In the second part of this talk I will present some recent data (German and D’Imperio, submitted) on the phonological realization of information focus and its scope in French wh-questions. In contrast with stress-accent languages such as Italian, Spanish or English, French does not appear to signal focus through pitch accent assignment, rather it appears to mainly exploit prosodic edge marking for the same purposes. The fact that prosodic phrasing is highly sensitive to focus structure is not only true for French, but also for pitch accent languages such as Japanese and Basque (see Gussenhoven 2004 for a discussion), as well as for stress-accent languages (Beckmann & Pierrrehumbert 1986). Different languages use different methods to signal phrasing, like segmental or tonal sandhi rules (cf. Shih 1990 for Chinese) as well as intonational structure.

A previous analysis (Féry 2001) has proposed that French largely exploits phrasing in order to signal focus, and that narrow and contrastive focus “lead to an initial boundary tone, usually high”. Here I shall attempt to build upon Féry’s insight by showing that, while phrasing is one of the strategies that French adopts in order to signal focus, phrasing cues are different when either the left or the right edge of the focal domain are taken into account. Right edge marking can either reflect an Intonation Phrase (IP) or an intermediate phrase (ip) break, signalled through a H-tone (see Fig. 1). The role of syntax/prosody alignment constraints on the placement of phrase boundaries has been shown for various languages (Selkirk, 2000, Truckenbrodt 1999, Feldhausen 2008). Our assumption is that the emergence of an ip in French is not simply linked to a specific focus or marked syntactic structure so that an ip boundary can occur within broad focus utterances when the syntactic structure allows it (Michelas and D’Imperio to appear, D’Imperio and Michelas, submitted).

While the intermediate phrase break is generally syntax driven, the intonation phrase break seems to be obligatory when the focus domain is narrow (i.e. restricted to part of a DP, such as an
adjective or a demonstrative). The other novel claim is that focally induced phrasing is marked through the presence of a left edge (an initial rise, or LHi, in the sense of Jun and Fougeron, 2000, as on the initial syllable of marron in Fig.1, upper panel), while syntactically induced phrasing does not appear to require it. Also, the initial rise is claimed to be a phonologically and phonetically different phenomenon than the final rise (contra Féry 2001). While the phonetic implementation of the initial rise is characterized by a marked lengthening of the word onset consonant as well as by higher intensity (see Portes and D’Imperio, 2008, submitted), the final rise is mainly marked by rhyme lengthening and its placement is restricted to the last strong syllable of the phrase. Moreover, strategies such as dephrasing and/or downstepping of material that is either prefocal or internal to a large focal domain seem to be exploited in order to enhance the focus contrast (see different tonal realizations of valise “suitcase” in Fig. 1). Finally, the placement of an initial boundary does not appear to be restricted to the left edge of a Maximal Projection, but can occur to the left of an argument within a complex syntactic constituent when focus is restricted to a lexical item.

**Figures**

Fig. 1 F0 curve and spectrogram for the sentence *Mais à qui est-ce qu’Amélie a vendu la valise marron dans la rue Mignet?* “But to whom has Amelie sold the brown suitcase on Mignet Street?” uttered with either narrow focus on the adjective (upper), with NP focus (medial) or with VP focus (lower). Note the dephrasing of the noun valise in the last example.
Meanings, shades of meanings and prototypes of intonational categories

Barbara Gili Fivela
Università del Salento & CRIL - Lecce
barbara.gili@ateneounile.it

In the past years various studies have focused on the perceptual study of intonational categories, with a specific interest in the categorical perception of intonational events (Niebuhr & Kohler, 2004; Schneider et al.; 2006; Vanrell, 2006). Most of these investigations have been relying on the categorical perception paradigm (Lieberman et al., 1957) for checking, by means of both identification and discrimination tests, whether intonational contrasts could indeed be perceived through categorical perception. Nevertheless the paradigm has been criticized by some authors because of the problematic results in the perception of intonational events (Massaro & Cohen, 1983; Massaro, 1998; Niebuhr & Kohler, 2004; Schneider et al., 2006), along the lines of what has already been observed in the perception of vowel sounds (Schneider et al., 2006). However the inconsistency in the results obtained in relation to categorical perception of intonational events is quite puzzling and deserves some considerations. On the one hand, different results are found in the literature for different tonal events (e.g., peak and valley contours in German – see Niebuhr & Kohler, 2004) and ascribed to the specific properties in the alignment of the events to segments. This questions the robustness of the methodology, as some authors observed, and suggests that “the perception of categories may be more important than the categorical perception of events” (Kohler, 2006). On the other hand, the situation may be even more puzzling in that the variability in results relates to similar patterns (e.g., a rising on the final syllable), showing similar contrasting functions in different languages (e.g. statement vs question function in Dutch and German; see Remijsen & van Heuven (1999) and Schneider et al. (2006) respectively). This suggests a complex situation in which the tonal event acoustic characteristic may not be the only relevant factor.

On the basis of the results of categorical perception tests run for the variety of Italian spoken in Pisa, it will be argued that the puzzling results may be due to both the way a (variety of a) language exploits the acoustic space and the lack of discreteness in the tonal categories investigated (Gili Fivela, 2008). Indeed the way a (variety of a) language exploits the acoustic space is taken to play a role in favoring/allowing CP; moreover, the functions and meanings related to intonation may be not discrete, although they may still be considered as linguistic (e.g. focus). In Xu’s (2006) words, the functions may overlap in case they “can both remain operative despite the overlap”; here it will be argued that the presence of shades of meanings interfere with the presence of categorical perception (Gili Fivela, 2008). Thus a different paradigm, among those focused on perception, will be considered for testing tonal events in Pisa Italian. The idea is that the peculiar situation related to intonation may be better analyzed with reference to the presence of prototypes of intonational categories. In this respect, data on the Perceptual Magnet Effect (Kuhl, 1991; Schneider et al., 2006) in Pisa Italian will be discussed, with the aim of checking whether the paradigm offers a better account for the presence of tonal events that are not categorically perceived although they express different functions and meanings; a second goal is trying to gain a more detailed view of the perceptual effects of the shades of meanings expressed by means of intonation.
References


On linguistic and paralinguistic meanings of intonation
Carlos Gussenhoven
Radboud University of Nijmegen
c.gussenhoven@let.ru.nl

If we ignore the role of intonation in marking constituency, the meaning of intonational structure concerns the way the proposition relates to the discourse model. That is, it answers the question what the listener is to do with the propositional meaning: evaluate it for its truth or relevance, take it as read, assume he already knew or knew part of it, consider it important, consider it routine, take it as a historical update, take it as a description of the existing world, emphasize the limits of its meaning, etc. These meanings can be classed as ‘information structure’ in a wide sense of the word. I can illustrate these meanings, but cannot give a definition such that meanings which ought to fall outside it are excluded (like ‘warning’, ‘sarcasm’, ‘promise’), and which are to be understood as pragmatic inferences of the bare intonational meanings. Another unanswered question is why intonational meaning should be restricted the way it is, and why it couldn’t mean ‘sarcasm’ or ‘evidentiality’ or ‘promise’, or to what extent the view expressed here is mistaken, and based on data from too few languages. Also, given the wider palette of meanings expressed in the phonetic implementation, the restrictive semantics of structural intonation is not self-evident.

In the phonetic implementation, speakers express meanings that derive from ‘biological codes’, meaning dimensions based on aspects of the production process of pitch variation, following John Ohala’s work on the Frequency Code. The meanings are metaphors of the conditions that are responsible for variation in the vocal fold vibration rate which are by-products of the speech production process. The Frequency Code accounts for meanings that are derived from inferences from the small size of a creature producing high pitch, because smaller vocal folds produce higher rates of vibration. The Effort Code accounts for meanings that are derived from inferences from canonical, careful, pronunciation. The Respiratory Code, to borrow Francis Nolan’s term, accounts for meanings that are inferences from locations in the breath group, which has higher pitch at the beginning than at the end. Yi Xu’s recently proposed Vomit Code accounts for meanings that are inferences from the gagging posture of the tongue and possibly other physical features of vomiting. None of the meanings require the presence of the physiological circumstances that lie at their basis; but their effects are used in partly conventionalized improvisations on the metaphorical meanings. Substitute features may be used, such as when late peaks are used to signal the meanings of high peaks, and speech communities may vary in the way the biological codes are employed. I will illustrate these views with results from experiments with Aoju Chen reported earlier.
Prosody and context management in interactive discourse: 
a study of two different interaction scenarios.

Jill House
University College London
j.house@ucl.ac.uk

In the course of an interaction, talkers engage in a process of building and updating the context within which their contributions should be understood. This context is multi-layered; it involves, for example, the linguistic content of the message, any paralinguistic information, and the process of constructing the interactive discourse itself (House 2007). In a sense we are dealing with a range of contexts, all simultaneously under construction, but not necessarily equally salient at any given point. In Relevance-Theoretic terms, the 'context' will consist of the set of assumptions that are salient at any given time – the cognitive environment – and comprehension will involve assessing new information against existing assumptions and updating the active context accordingly (Sperber & Wilson 1995). Speaker and hearer together take responsibility for updating the cognitive environment: the speaker must use whatever communicative resources are available to allow the hearer to access the intended new set of assumptions, while the hearer's task involves accessing the relevant cognitive context within which to evaluate and interpret the speaker's contribution. At the same time, participants in the interaction will be constructing the discourse itself, negotiating and managing their turn-taking in a more or less orderly way (Couper-Kuhlen & Selting 1996).

Prosodic resources play a crucial and essentially procedural role in shaping talk so as to guide access to relevant assumptions, whether these be linguistic, paralinguistic or oriented towards constructing the interaction itself. However, the same resources are not uniquely associated with any particular context orientation, so when faced with potentially ambiguous prosodic cues, the hearer will access whichever assumptions are least costly in terms of processing effort.

This talk will explore the prosodic choices made in excerpts from two very different discourse scenarios. The first scenario looks at excerpts from an intrinsically co-operative type of discourse, the Map Task (Shobbrook & House 2003; House 2007), where participants work towards a common goal. The second scenario is somewhat combative, involving excerpts from the BBC TV show The Apprentice (2009), in which participants are competing against each other in the board room to save their place in the programme, thus pursuing an individual rather than a collective goal. Using evidence from the interactions themselves, the contribution of specified prosodic phenomena (tonal choice, pitch range; tempo) to context management and the construction of the interaction will be assessed.

References

One of the key questions in language studies is how the speech signal is segmented into word and phrase-sized units. A way to address the problem is to examine the processing of utterances with temporary ambiguities. Prosody is known to constrain both lexical access and syntactic analysis in ambiguous sentences (Kjelgaard & Speer 1999, Christophe et al. 2004, Millotte et al. 2007, 2008, Dilley & McAulley 2008). Although there is general agreement on the disambiguating role played by intonational phrase (IP) boundaries, the role of lower phrase boundaries, as well as of word boundaries, seems to be still an open issue, in different studies and cross-linguistically (e.g. Price et. al 1991, Kjelgaard & Speer 1999, Christophe 2003, Li & Yang 2008). In this paper, we address the role of prosodic structure in the processing of both lexical and syntactic ambiguities in European Portuguese (EP), specifically we examine the effects of the full range of prosodic boundaries reported for EP in previous literature: prosodic word (PW), prosodic word group (PWG), phonological phrase (PhP) and IP (Frota 2000,2008; Vigário 2003, 2009). To our knowledge, this is the first study addressing this issue for EP, and one of the few that includes a full set of prosodic structure contrasts.

We tested the effect of prosodic structure by means of two experiments: a completion task (Exp.1) and a word detection task (Exp.2). Pairs of homophonous words were used to create utterances with a local ambiguity, which were read by an expert speaker and a naïve speaker. Utterances were cut right after the end of the ambiguous word, and participants listened to the cut utterances in two blocks so that each member of an ambiguous pair only appears in one of the blocks. In Exp. 1, participants were asked to write the listened sentence and complete it as they saw fit in a response sheet, after listening to it as many times as they wished; in Exp. 2, participants had to detect abstract lexical entries for the ambiguous word (along the lines of Millotte et al. 2008). We will report on preliminary results of Exp.1, on the basis of data from 12 educated native speakers of EP (6 for each block) aged between 21 and 37 years old (396 responses=33 pairs X 12). The experimental conditions tested are shown in Table 1.

The descriptions of the prosodic structure of EP define a PW signaled by a constellation of cues (segmental and prominence cues), as well as a PWG also marked by segmental and prominence facts. By contrast, the PhP is argued to be weakly marked, although it plays a relevant role in rhythmic phenomena and pitch accent distribution. The IP is marked by all sort of cues: sandhi, lengthening, pitch accent and boundary tone (Frota 2000, Vigário 2003, 2008). If these cues are exploited in processing, we predict that a stronger effect should be expected for IP and the word constituents than for the PhP. However, our results run against this prediction. As shown in Fig.1, listeners clearly distinguish the two kinds of word boundaries from the PhP, the PW boundary from the IP, and the PhP from the IP, thus suggesting a disambiguating effect of prosody at all these levels. ANOVAs have shown that the effect for each pair of prosodic structure comparison is significant (p<.01). By contrast, the boundary between 2 PWs within a PWG is not distinguished from that between 2 PWs within a PhP, and the within PW condition is not different from the PW boundary condition. In the former pair, there is a strong response bias towards a [PW PW]PWG pattern; in the latter case, responses vary in both conditions (but with across items effects). There is no significant main effect of listeners (f(11,246)=.44; p<.9343). These results support a disambiguating role of the PhP in EP, and not just the IP; they also show that PW boundaries are hard to detect as different from within PW contexts, a result similar to that reported for Chinese in Li & Yang (2008), on the basis of EEG measures, and that calls for further and deeper research.
Figures and Tables

Table 1. Experimental conditions

<table>
<thead>
<tr>
<th>Prosodic structure pairs</th>
<th>Examples of test items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within PW / PW boundary</td>
<td>Gosto da [pintadela]PW que deste ao armário.</td>
</tr>
<tr>
<td>PW boundary / Within PWG</td>
<td>O miúdo foi [(mal]PW [criado])PW com os avós.</td>
</tr>
<tr>
<td>Within PWG / PhP boundary</td>
<td>A estagiária [surda-muda]PW desfez os preconceitos alheios.</td>
</tr>
</tbody>
</table>
| Within PWG / PhP boundary              | A estagiária surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surda surd...
The role of duration and tonal scaling as complementary cues in distinguishing presumptive from neutral yes-no questions

Verònica Crespo-Sendra,1 María del Mar Vanrell,1 Pilar Prieto1
1U. Pompeu Fabra, 1ICREA-UPF
veronica.crespo@uab.cat, mariadelmar.vanrell@uab.cat, pilar.prieto@uab.cat

Previous studies have dealt with the phonological contrast that seems to exist between echo questions that convey incredulity and other yes-no question types such as neutral (Lee et al. 2008) or uncertainty yes-no questions (Hirschberg & Ward 1991). Hirschberg & Ward (1991), in their study about the difference between utterances that convey incredulity and uncertainty, confirm the existence of distinct pragmatic interpretations when the only change between two utterances is the pitch range. Lee et al. 2008 analyze the differences between neutral absolute and presumptive absolute interrogatives and conclude that the presumptive questions have an even wider pitch range and an expanded range in the nuclear peak relative to the pitch range and nuclear peak height of the neutral question; these differences characterize the pragmatic contrast.

In Catalan, neutral and presumptive yes-no questions have been characterized in the same way, that is, the first stressed syllable is realized with a falling tone followed by a rising movement, and the last stressed syllable is produced with a low tone followed by a sharp rise (Prieto 2002, Prieto and Rigau 2007). According to these studies what distinguishes these two types of yes-no questions is the pitch range. In the case of presumptive questions the pitch range is much broader.

In this paper we will investigate the contribution of F0 (related to the boundary tones) and the duration of the last syllable in the utterance (which was observed to be a potential important cue in presumptive contours in Prieto and Cabré, 2007) in distinguishing these two question types. To this end, a standard identification task was applied to this contrast. In addition, Reaction Time (RT) measurements were taken since some researchers have proposed the RT approach as a good alternative to the discrimination task in testing the hypothetical discreetness of a contrast (Pisoni and Tash 1974, Chen 2003). The materials for the identification task consisted of three continua in which three different parameters were manipulated from neutral to presumptive interrogative: duration, tonal scaling and a combination of both prosodic features. Thus, in the duration continuum the duration of the last syllable was manipulated from 190 ms to 240 ms in 10 steps of 5.5 ms each; in the tonal scaling continuum we manipulated the tonal scaling of the boundary tones from a high to an extra.high pitch in 10 steps of 12.5 Hz each; and finally a combination of duration and tonal scaling was manipulated each time in order to create the third continuum.

Pilot results with 5 listeners show that manipulation of duration has no effect in distinguishing presumptive from neutral yes-no questions (see Fig. 1). On the contrary, tonal scaling has an important effect (see Fig. 2) since the results show that speakers are able to change the interpretation from neutral to presumptive yes-no questions only by increasing the pitch height of HH% boundary tones. However, the RT measurements do not support a categorical distinction, since we cannot identify any clear peak at the crossover point between the two categories. Interestingly, it is the combination of duration and tonal scaling that display the nearer categorical results: there is an abrupt change in the response rate between stimulus 2 and 3 that corresponds to a plateau in the RT measurements. On the basis of these results, we may conclude that it is the combination of cues rather than any one cue taken independently that Catalan listeners use when distinguishing neutral from presumptive yes-no questions.


Prosody in German Sign Language
Annika Herrmann
University of Frankfurt/Main
herrmann@lingua.uni-frankfurt.de

In this presentation, I argue that German Sign Language (DGS) exhibits a prosodic system equivalent to spoken languages. This first systematic investigation of prosody in DGS, based on annotated video data, provides an interesting testing ground for the hypothesis that some intonational cues in sign languages have general meanings and are componential in nature. The systematic use of prosodic cues in DGS supports the analysis of prosody as a universal property of language.

As prosody is concerned with rhythm, pitch, and intonational aspects of language, it is interesting that research on languages in the visual-manual modality has revealed that prosody does not necessarily need to be linked to vocal articulation. The system of prosody in sign languages offers interesting results with regard to modality independent notions of language constitution. As an interface phenomenon interacting with different levels of grammar, prosody in sign languages requires the same analytical devices as in spoken language theories.

Sign languages segment and structure the discourse analogously to the prosodic hierarchy found in spoken languages (Nespor & Vogel 1986). Equivalent units for prosodic words, phonological phrases (PP), and intonational phrases (IP), for instance, have been attested in various sign languages such as American Sign Language, Israeli Sign Language, and the Sign Languages of the Netherlands (cf. Wilbur 2000, Brentari & Crossley 2002, Nicodemus 2009, Nespor & Sandler 1999, Crasborn & van der Kooij 2004), but we lack systematic analyses on prosody in DGS.

In sign languages, meaning is conveyed by various articulators: the hands (manual) and others such as the upper body, the head, and the face (nonmanual). Nonmanual features (NMFs) like eye brow movements and head nods are systematically used for prosodic marking (Pfau & Quer in press, Pfau 2006). Some NMFs have been analyzed as having inherent meanings that can also be combined to derive complex meanings (Dachkovsky & Sandler 2008). As the same cues were also found in DGS, the hypothesis has been tested for DGS resulting in supportive evidence.

I investigated manual and nonmanual prosodic cues on the basis of video data of deaf native signers that participated in an elicitation study. The results were annotated by ELAN, a tool for linguistic video annotation and analysis (MPI Nijmegen). Various cues interactively structure signed utterances in DGS. I will present the prosodic system found in DGS and concentrate on the marking of PPs and IPs. The distinction between rhythm, prominence, and intonation in DGS will be clarified by examples of the respective cues and the difference between domain and edge markers. Domain markers may spread along prosodic constituents and thus are comparable to intonational aspects of prosody, whereas edge markers occur punctually to indicate phrase boundaries and segment the discourse into rhythmic units (see figures below).

In the discussion about the prosodic nature and status of cues that fulfill rhythmic, intonational, and prominence marking functions, many factors such as syntactic constituency and the syntax - prosody relation (Selkirk 1984), the frequency and systematic use of cues, and particularly the interaction of various features play a crucial role. Boundary markings such as eye blinks, head nods, frozen signs, etc. may be seen as equivalents to phrase final lengthening and boundary tones in spoken languages, for instance, and give further support to the analysis of prosody as a universal concept.

The findings from DGS give insight into how visual languages realize phrase structuring, phrase boundaries, and prosodic marking and to what extent they use the system consistently.
Figures

Figure 1: Template of the prosodic marking of DGS utterances

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>b b b b ln ln ln</td>
<td>eye blinks head nods</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>b b b b ln ln b</td>
<td>eye blinks head nods eyebrows body facial expr. mouth</td>
</tr>
</tbody>
</table>

“A man sits in the living room reading a book. A woman comes along and brings him coffee. The man lazily leans back while reading. Suddenly the doorbell flashes.”

Figure 2: Prosodic markings in the above DGS template (Figure 1)

2.1 Eye blink and head nod after SIT  2.2. Same nonmanuals on BE-LAZY and READ

2.3 Nonmanuals on SUDDENLY  2.4 Nonmanuals on FLASH

References


Poster presentations
In this paper, we address the issue of the informational structure and prosodic system interface in Tuscan Italian (TI). We propose a model of how the informational properties of focus (background) along with (left and right) topics, and “Givenness” (Schwarzschild 1999) feed the PA association mechanism. We performed a production experiment using read speech (4 speakers x 46 short scripts x 5-8 repetitions). The scripts were designed in order to induce different types of focus structure in combination with left and right dislocated Topics in different linear orders. Such a design allowed us to factor out the linear order in which focus appears and to evaluate its impact on other phrases endowed with discourse properties, such as dislocated topics. The corpus has been analyzed and ToBI-transcribed. In our corpus, “Given” information occurring before focus is systemically stressed, phrased and accented. This result fully supports the conclusion that Italian fails to deaccent given information (Ladd, 1996; Swerts et al., 2002). In postfocal contexts the pitch contour is low and flat, irrespectively of type, number and length of the constituents.

We show that the lack of pitch movements on postfocal phrases in TI cannot be accounted for neither in terms of “deaccenting/destressing given” (Féry & Samek-Lodovici 2004), nor in terms of extametricality of postfocal material, nor as the result of tonal unspecification. Indeed, we argue that postfocal elements, in spite of the apparent lack of PAs, are metrically phrased and phrasal metrical prominences are assigned. This conclusion is supported by Bocci (2009) which shows that despite the lack of pitch movements, the effects of phrasal prominences in postfocal contexts are detectable as they involve longer vowel durations, spectral slope differences and more extreme formant values.

We claim that postfocal material is visible to intonation: it is associated with a L* PA. A consequence of such a proposal is that the difference among postfocal prominences in TI, Southern varieties of Italian (Grice et al. 2005) and European Portuguese (Frota 2000) is reduced to a typological one: while the latter associate with special downstepped PAs, TI associates with L*. We propose that L* in TI is a marker exclusively dedicated to define the right edge of the Focus phrase by marking the material to its right as non-focal. L* is incompatible with a focus interpretation. We argue that L* is neither the prosodic correlate of the background, nor of Vallduvi’s (1993) “tail”. We propose that L* distribution is ruled by a Focus Defining Rule which sketchily requires that the focus phrase is associated with the rightmost PA able to express focus (where L* in TI is unable to express focus, as !H+L* is in EP). For instance, minimal pairs of clauses with in situ and ex situ focus (elicited by the same context) share presumably the same focus/background partition and the focus phrases bear invariantly the same PA. However, the prosodic properties of the background are radically different in the two cases: prefocal material bears PAs different form L*, while FDR forces postfocal material to associate with L* (see fig. 1, 2). Building on Calabrese (1992), we assume that Italian allows only one focus per utterance. We show that the domain of FDR in Italian is the Utterance as a follow-on from the uniqueness requirement and that FDR outranks any other PA association. Parenthetical and (certain) adjunct clauses, CLDED and RDED topics are phrased into independent intonational phrases metrically headed. If they precede focus, they associate with PAs different from L*, while, if they occur after focus, they invariably associate with L* (see fig. 3, 4).

According to our proposal, L* is not a correlate of any informational property, but it is inserted to define the right edge of the focus phrase, overriding any other PA association. We argue that PA association cannot be exclusively driven by the informational properties of the element which a PA is associated with (contra Pierrehumbert & Hirschberg 1990), but that it is also ruled by structural requirements on the intonational representation.
### References


Early acquisition of form and meaning in Catalan and Spanish interrogatives

Joan Borràs-Comes,1 Verònica Crespo-Sendra,1 Maria del Mar Vanrell,1 Jill Thorson,6 Pilar Prieto1

1Parc Científic de Barcelona–UPF, 2U. Pompeu Fabra, 3Brown University, 4ICREA–UPF
jmborras@pcb.ub.cat, jcthorson@gmail.com, mariadelmar.vanrell@uab.cat, pilar.prieto@upf.edu, veronica.crespo@uab.cat

The link between melody and meaning has been a debated topic of past and recent research. The possible connection between the rise and fall of the fundamental frequency and the semantic and pragmatic intentions of the speaker has been proven to be a compelling one. Specifically, the choice of intonation contour for an interrogative utterance has been claimed to correlate to a variety of different pragmatic meanings of the speaker (for Spanish, see Navarro Tomás, 1968; Escandell, 1998; Sosa, 1999; Sosa, 2003). Thus, the distinct contours and nuclear configuration patterns of interrogatives signal different interpretations of the utterance. According to Escandell (1998; see also Sosa 1999, 2003), Spanish interrogatives fall into one of three different prosodic patterns depending on the pragmatic intention. The first type is the falling-rising ending that conveys an interpretation of a thought that would be relevant to someone if true (ex: L+H* L%). The second type is a rising-falling ending that imposes a further attribution to another individual (ex: L+H* LL%), while the third type is a final rising contour that imposes this attribution to the speaker (ex: L+H* HH%) (Escandell, 1998). Therefore, the nuclear configuration patterns reveal different intentions of the speaker.

For this study, we will investigate this link between intonation and meaning in child-directed speech and how this is reflected in the prosodic and pragmatic productions of interrogatives by Catalan and Spanish speaking children. Previously, it has been found that children produce several types of interrogatives during early speech development and now we would like to explore the correspondence between intonational form and meaning (Prieto et al. 2008). Thus, the goals of this paper are: a) to analyze prosodically the interrogative forms produced by 2 Catalan- and 2 Spanish-acquiring children between the ages of 17 and 28 months b) to relate the results of this analysis to the development of the capacity of using interrogative forms in a proper context based on the “speech act” concept and c) to investigate the influence of the child-directed speech forms on the productions and intentions of the children (Austin, 1962; Searle, 1969; Grice, 1957, 1975; Holzman, 1972).

For this study, 301 yes-no questions for Catalan-acquiring children and 319 for the Spanish-acquiring children were first analyzed pragmatically in Phon and then independently analyzed prosodically in the AM framework in Praat (Rose et al., 2006; Pierrehumbert 1980; Boersma & Weenink 2009). Because formal language development in children seems to emerge from pre-linguistic social, cognitive and perceptual interactions with adults, interest has increased in the development of the notion “speech act” as the basic unit of linguistic communication (Holzman, 1972; Dore, 1975; Greenfield et al., 1976; Rodga et al. 1977; Bates et al., 1977). Therefore we propose to analyze the utterances of the children following the notion of identifying the underlying pragmatic form (i.e., query, request, test, offer, check, object, align). Our pilot results demonstrate that all children have performed some instance of questioning before the two-word period and that their productions do in fact reflect the adult inventory pattern (as in Lleó & Rakow, 2008), producing both falling (H+L* L%) and rising contours (L* HH% and L+H* HH%). Interestingly, both Catalan and Spanish children start to produce the rising contour L+H* HH% first. This nuclear configuration, namely L+H* HH% opposed to the L* HH% configuration, has been characterized as suggesting a “helpfulness” meaning, especially in child-directed speech (Escandell, 1998; Escandell, 1999; Gussenhoven, 2002). In the paper, we will present the results from relating the different types of nuclear configurations presented above to the different intentions that can underlie an interrogative form.
References


In language acquisition, a construction of particular importance is the basic transitive construction, prototypically used to indicate an agent acting on a patient, as in “The Flomer weefed the Miemel”. To interpret such transitive constructions one needs to understand and to distinguish the different roles of participants in such an event, i.e. to understand the grammatical conventions used to mark the participant roles in the particular language being learned. In most languages the listener has multiple, sometimes redundant cues (e.g., word order, case marking, or animacy) in order to distinguish the participant roles in an utterance. These cues are acquired step by step. For German, a language with case marking and relatively free word order, Dittmar et.al. (2008) found that two year olds only understood sentences, where several cues supported each other. At the age of five, children were able to use word order by itself but not case marking and only 7-year-olds behaved like adults by relying on case marking over word order when these two cues conflicted (e.g. “Den (+accusative) Löwen wieft der (+nominative) Hund” – “The (+accusative) lion is weefing the (+nominative) dog”).

However, most studies examining children’s understanding of transitive constructions focus on the morpho-syntactic properties of sentences and ignore an additional cue: prosody. But, it has been established that listeners’ interpretation of ambiguous sentences is guided by prosodic phrase boundaries and their realization (e.g. boundary tones, final lengthening, pausing) as well as the presence or absence of pitch accents. Grice, Weber & Crocker (2006) found that adult-listeners use prosodic information in the interpretation of ambiguous SVO and OVS sentences when no clear morphological information is available. The prosodic information was integrated rapidly enough to affect listeners’ interpretation of grammatical function and the assignment of participant roles before disambiguating case information was available.

In the current study we investigate whether or not German learning children at the age of 5 use prosody for the assignment of participant roles in order to distinguish their semantic roles, as has been found for adults. Using a video-pointing task, we embedded transitive OVS utterances in a natural context and presented these utterances as either clearly case marked (e.g. “Den (+accusative) Hund wieft der (+nominative) Hase”) or ambiguous (e.g. “Die (+accusative) Katze wieft die (+nominative) Kuh”). In order to examine the specific role of prosody for children in resolving the semantic functions of the participants, the intonational realization of these constructions was either flat and deaccented or, to support the syntactic marking of the utterance, characterized by a strong, contrastive L+H* pitch accent on the first Nominal phrase.

The results show that the prosodic cue has a main effect for children for the assignment of participant roles in transitive OVS utterances (F(1,15)=5.8, p= 0.029). Children were better in judging the correct agent acting on the correct patient when this was clearly marked by intonation compared to realizations without any intonation. Even when no clear case marking was available, children understood participant roles significantly better by using the prosodic cue (p=0.009) (see Figure 1). These findings show that, when reliable cues contradict each other, children at the age of 5 are still able to understand the semantic roles in transitive OVS sentences when appropriate intonation is available. This suggests that children use prosodic cues to understand the grammatical conventions of a particular language from early on. Thus, prosody is an important feature for children in the language acquisition process.
Figures

Figure 1:

References

This study presents a production experiment investigating (i) the tonal and durational means of encoding focus and givenness, and (ii) the interaction of focus and downstep in Akan, a tone language that belongs to the Kwa branch of the Niger-Kongo family spoken by 5 mio. people in Ghana. According to Kobele & Torrence (2006) the expression of pragmatic prominence like focus is realized by means of a syntactic cleft construction that is additionally marked with the focus particle “na”. Prosodically, Boadi (1974) reports a raising of high and low tones in this position. This kind of interaction of tone and intonation has been reported for other geographically unrelated tone languages as well (e.g. Mandarin Chinese, Xu 1999; Ewe, Fiedler & Jannedy, submitted; Kammu, Karlsson et al. 2007), though not for tone languages in general (e.g. Kügler & Skopeteas 2006, 2007 for Yucatec Maya).

We examined the information structural effects recording 11 speakers uttering sentences with in-situ (1) and ex-situ focus constructions (2). The ex-situ structures were chosen to investigate whether the reported tonal raising effect due to focus is triggered by prominence or whether it is solely positional in nature. The in-situ structures were chosen to look for particular prosodic effects when no other means of prominence marking are present. Target words were controlled for tone using a word associated with low tones (1), one with a high tone (3), and one which we assumed to be lexically unspecified for tone (4). These words were embedded in carrier sentences keeping the tonal context around the target words constant. The test sentences were preceded by different context questions that trigger a different information status on the target words each: Target words were either in broad, narrow (1), and contrastive focus, or either pre-focally (4) or post-focally (3) given. We found that duration plays no vital role in the prosodic expression of focus. However, the information status of a constituent has an effect on the tonal realization: Contrary to Boadi (1974) and contrary to a general tendency of languages to raise F0 on focused constituents, the lexically low and high target tones in the ex-situ construction are lowered under focus. This effect shows up also in-situ for the high tone and the toneless target word. Whereas the low tone target is raised due to focus. Givenness seems not to be marked consistently.

The interaction of focus and downstep may be a further indicator of a prosodic means to express pragmatic prominence when focus blocks the downstepping of tones as it is the case for many other languages, e.g. German (Féry & Kügler 2008), Japanese (Ishihara, submitted). This blocking may result in tonal upstep on the focused constituent. Downstep in Akan requires an underlying H-L-H tone sequence and may be automatic (5) (downstep of the second H tone, H-L-H) or non-automatic (6) (downstep of the H tone prior to a floating L tone, H-LH). To investigate the effect of focus on downstep we recorded sentences containing four different constructions where the tonal environment for downstep is met. These target phrases were embedded in different contexts that trigger wide and narrow focus sentences, and sentences with contrastive focus either on the first (5) or second target word (6). Our data do not show a blocking of downstep due to focus, i.e. downstep occurs in all conditions independent of the information status.

Taken together the results indicate that pragmatic meaning appears to be realized prosodically independent of syntactic or morphologic focus marking. Further, downstep is realized independent of focus indicating its special grammatical role in Akan which cannot be blocked for pragmatic reasons.
Examples
(Target word in bold)

(1) Q: Hwan na Agyeman boa anopa yi?
   Q FM Agyeman help-post.3sg morning this
   When did Agyeman help this morning?
   A: Agyeman bɔ̀nɔ̀ Addo ɔ̀nɔpɔ̀ yi.
   Agyeman help-post.3.sg Addo morning this
   Agyeman helped Addo this morning.

(2) A: Addo na Agyeman boa anopa yi.
   Addo FM Agyeman help-post.3.sg morning this
   It is Addo that Agyeman helped this morning.

(3) Q: Anum dii amango anopa yi?
   Anum eat-post.3.sg mango morning this
   Did Anum eat a mango this morning?
   A: Daabil Anum toɔ amango ɔŋɔpɔ yi.
   Nel Anum buy-post.3.sg mango morning this
   Nel Anum bought a mango this morning.

(4) Q: Gyata wee Ananse enora?
   Lion eat-post.3.sg spider yesterday
   Did a lion eat Ananse yesterday?
   A: Daabil Gyata wee Ananse ɔŋɔpɔ yi.
   Nel lion eat-post.3.sg Ananse morning this
   Nel A lion ate Ananse this morning.

(5) Q: Anum rebɔ Agyeman Papa breco?
   Anum beat-prog 3sg. Agyeman father softly
   Is Anum beating Agyeman’s father softly?
   A: Daabil Anum rebo Kɔfi papa breco.
   Nel Anum beat-prog 3sg. Kɔfi father softly
   Nel Anum is beating Kɔfi’s father softly.

(6) Q: eye Kɔfi dua na eretuttu abunu?
   be-pres 3sg. Kɔfi plant FM against fever
   Is Kɔfi’s plant helping against fever?
   A: Daabil Kɔfi nasa no rebɔ atutu abunu.
   Nel Kɔfi water help-prog 3sg. against fever
   Nel Kɔfi’s water is helping against fever.

References

Ishihara, S. (submitted) Japanese Downstep Revisited. NLLT.
Questions headed by the particle “que” in the Spanish spoken in Lleida: a prosodic bilingualism interference
Eugenio Martínez Celdrán, Ana M. Fernández Planas, Lourdes Romera Barrios & Josefina Carrera Sabaté
University of Barcelona, Laboratory of Phonetics
martinezceldran@ub.edu, anamariafernandez@ub.edu, lromera@ub.edu, jcarrera@ub.edu

The use of questions headed by the unaccented interrogative particle que (‘that’) is very common in Catalan, and the possibility of using such introductory forms is banned in Spanish. Nevertheless, these questions can be used in the Spanish spoken by bilinguals from Lleida as well as those from other Catalan areas. This fact shows a clear interference effect on Spanish prosody due to bilingualism, especially in questions headed by the unstressed interrogative particle que (‘that’).

The AMPER Project seeks to collect and describe prosodic samples of Romance languages for the elaboration of a large-scale multimedia atlas. In Catalonia, the development of the project for the two official languages, Spanish and Catalan, is known as AMPERCAT (Fernández Planas, 2005; Fernández planas and Martínez Celdrán, 2003, 2007). Within this framework, the present paper examines Spanish interrogatives sentences and declaratives sentences, the former with two forms: interrogatives with and without the unstressed conjunction que in sentence initial position. The sentences were produced by a female informant whose first language is Spanish but who shows noticeable influence of Catalan having been born in Lleida, a Catalan-Spanish bilingual area.

The analyses were performed on a corpus of 567 sentences which focused on the initial and final values of F0 vowels from each sentence. A similar intonation contour to that of the Catalan of Lleida has been observed for Spanish declarative sentences and interrogatives with and without the unaccented particle “que”. Even if the latter interrogatives is not to be found in the descriptions of Spanish intonation, it is commonly observed in Spanish speakers residing in Catalonia, as a result of the influence of Catalan intonation contours, which combine both interrogatives (¿La guitarra se toca con paciencia? ‘Is the guitar played with patience?’ vs. ¿Que se toca con paciencia la guitarra? ‘That is placed with patience the guitar?’). Data can also be described through the AM theory (Autosegmental-Metrical Theory, Pierrehumbert 1980 and Prieto 2002) in the new formulation of SpToBI (Estebas and Prieto 2008).

As a general conclusion, and accordingly with other studies (Prieto, 2002; Romera et al. 2007, 2008), the passive knowledge of the prosody of any non-native contact language, in this case, Catalan, appears to exert a strong influence in modifying or determining the prosody of the speaker’s mother tongue, here Spanish. This is shown by the consideration of Spanish interrogatives formed by inserting initial unstressed que –hence closely resembling Catalan interrogatives (Carrera et al., 2004; Fernández Planas et al., 2006, 2007; Martínez Celdrán et al., 2005a, 2005b, Romera et al., 2009)- as widely accepted “standard” interrogative sentences.
References


Endnote: This study was supported by Grant HUM2006-05238/FILO from the SGPI of MICINN.
The effects of prosody on the comprehension of novel noun-noun compounds has been somewhat neglected to date, despite its central role in linguistic communication (e.g., Bresnan, 1971; Ladd, 1996). Such conceptual combinations (e.g., cactus beetle) are highly polysemous, but most meanings either involve a linking relation (e.g., beetle-eats-cacti) or a property transfer (e.g., a prickly beetle).

The current study presents participants with novel compounds auditorially and, since prosodic emphasis affects activation of a word (Bock & Mazella, 1983; Cutler & Fodor, 1979), examines how changes in emphasis influences conceptual combination. All participants receive the same set of novel compounds but with emphasis on different words: modifier emphasis (e.g., OCTOPUS apartment), dual emphasis (e.g., OCTOPUS APARTMENT), or head emphasis (e.g., octopus APARTMENT).

Experiment 1

Thirty participants listened to 27 novel compounds (randomly interspersed with filler items) through headphones and pressed “yes” if they could think of an interpretation (then typed it onscreen) or “no” if they could not. Interpretations were coded blind to condition.

Means are given in Table 1. People generated more relation-based than property-based interpretations \[F(1, 28)=12.805, p=0.001\], but proportions showed no other effects or interaction. Response times showed an interaction between interpretation type and emphasis type \[F(2, 56)=5.354, p=0.007\]; planned comparisons showed that relation-based interpretations were faster than property-based interpretations for modifier emphasis \[t(28)=2.242, p=0.017\] and head emphasis \[t(28)=2.962, p=0.003\], but the opposite pattern was found for dual emphasis with property-based interpretations being faster than relation-based \[t(28)=1.706, p=0.0495\]. Simple effects analysis revealed that emphasis reliably affected property-based interpretation times, but not relation-based interpretations (p<0.05).

Experiment 2

This experiment removed the option of quickly rejecting a compound by asking participants to press the spacebar once they had thought of a meaning for the compound rather than offering a yes/no forced choice. Thirty new participants took part. Results replicated the previous experiment (see Table 1). People continued to produce more relation-based than property-based interpretations \[F(1, 28)=15.434, p=0.001\], with no other effects or interaction. Response times again showed an interaction \[F(2, 56)=3.565, p=0.035\]; in planned comparisons, property-based interpretations were again faster than relation-based interpretations for dual emphasis \[t(28)=2.159, p=0.04\], while the reverse trend was marginal for modifier emphasis \[t(28)=2.001, p=0.055\] and no difference in the head emphasis condition \[t=1\]. This demonstrates that dual-emphasis appears to facilitate the interpretation of property-based meanings relative to modifier- or head-emphasis. Once again, we observed that only property-based interpretations were affected by the emphasis manipulation (p<0.05).

Conclusions

Prosodic emphasis differentially affects the speed of interpretation of novel compounds, with emphasis affecting property-based interpretations but not relation-based interpretations. Such contrasting effects on distinct interpretation types are difficult to reconcile using single-process models of conceptual combination. Thus, we suggest that there are separate processes for property- and relation-based interpretations, in line with “dual process” accounts (e.g., Estes, 2003; Wisniewski, 1997) of conceptual combination.
**Figures and Tables**

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Emphasis Type</th>
<th>Proportion (%)</th>
<th>Response Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Property</td>
<td>Relation</td>
</tr>
<tr>
<td>Experiment 1</td>
<td>Modifier</td>
<td>36.9</td>
<td>57.6</td>
</tr>
<tr>
<td></td>
<td>Dual</td>
<td>34.6</td>
<td>61.6</td>
</tr>
<tr>
<td></td>
<td>Head</td>
<td>31.6</td>
<td>60.5</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>Modifier</td>
<td>37.4</td>
<td>52.1</td>
</tr>
<tr>
<td></td>
<td>Dual</td>
<td>27.8</td>
<td>57.1</td>
</tr>
<tr>
<td></td>
<td>Head</td>
<td>35.0</td>
<td>50.7</td>
</tr>
</tbody>
</table>

**References**


We investigated if priming experiments could be used to isolate meaningful prosodic features in interaction. Speakers in all age categories adapt their speech to their linguistic environment (Levelt & Kelter, 1982; Bock, 1986; Garrod & Anderson, 1987; Pickering & Branigan, 1999; Metzing & Brennan, 2003). On the phonetic and phonological level, they adjust to each others’ intensity and speech rate (Natale, 1975; Street et al., 1983), voice onset timing (Sancier & Fowler, 1997), global pitch range (Gregory & Hoyt, 1982; Gregory, 1986) and their phonetic realization of repeated words (Pardo, 2006; Delvaux & Soquet, 2007). While entrainment (also known as adaptation, alignment, convergence or mirroring) has been investigated rather thoroughly for adult speakers, less is known about younger age groups. Do children improve their entrainment abilities as their linguistic and social skills improve? Or do they adapt less as they grow older and develop their individual identities? On the basis of the Interactive Alignment Model (Pickering & Garrod, 2004), we might also expect that entrainment would be more likely to occur for linguistic features that are meaningful in a given interaction (i.e., receive a semantic ‘boost’), thus offering processing advantage to both interlocutors.

We report the results of three experiments in which 95 Dutch children in different age groups (4.5, 7-8 and 10-12 years) interacted with a pre-recorded computer voice. The goal of the experiments was to determine to what extent children adapt to various pitch properties of the interlocutor. We hypothesized that children would more likely adapt to pitch primes that were meaningful in the context compared to primes with no apparent functionality. In a mixed within-between-subject design, we systematically manipulated the global pitch properties of the voice (raised vs. lowered synthesis of the original recording) as well as boundary tones (H%, L%) in the nuclear phrase of each turn. The experimental task consisted of naming cards in an interactive computer game (fig. 1). In the course of the game, the manipulation of the global pitch had no apparent function but the boundary tones could be interpreted as either signaling continuation (H%) or finality (L%). In accordance with the predictions of the Interactive Alignment Model, we found that children were likely to adapt to the boundary tones with which they were primed, even when there was no detectable priming pattern (fig.2 and fig. 3), but did not adapt their global pitch. Furthermore, girls were more likely to entrain than boys and there was a decreasing trend in entrainment in the older age groups.
Fig. 1. Screen picture of the computer card game used in the experimental task. Children were taking turns with the computer in searching for three treasure chests hidden behind the cards in various card sets.

Fig. 2. Adapting to the low boundary tone prime in turn T (female child A, 4;1).

Fig. 3. Adapting to the high boundary tone prime in turn T+1 (female child A, 4;1).

References

Conditions for tonal curl and stød in the Central Swedish variety of Eskilstuna

Tomas Riad
Stockholm University
tomas.riad@nordiska.su.se

Background: The dialect of Eskilstuna (located 120 km to the west of Stockholm) belongs to the Central Swedish prosodic type. According to an earlier hypothesis (Author 2000), Eskilstuna constitutes a key dialect in understanding the relationship between Central Swedish and Dalabergslag Swedish, and in accounting for the phenomenon of generalized accent 2 which occurs in communities between Stockholm and Eskilstuna. Eskilstuna itself exhibits facultative curl/stød articulation of the prominence-boundary tones in certain contexts.

Prospect: A proper understanding of the relationship between these observed phenomena in Central Swedish can serve as basis for a historical reconstruction of a variety where stød occurs in grammaticalized form, namely Standard Danish.

This paper examines and document the general hypothesis that curl/stød arises in the realization of the prominence tone and the following boundary tone when these are compressed in one (preferably stressed) syllable. I take it that curl and stød are related articulations of a sharp HL% fall which may be uninterrupted (=curl, fig. 1) or interrupted (=stød, fig. 2). I look at emphatic articulations and compare them with unemphatic articulations in several speakers, the preliminary observation being that curl/stød correlates with emphasis, and that stød results from a relatively high degree of emphasis.

The realization of curl/stød also shows a leftward shift of the tonal contour. In Central Swedish, the prominence contour in focal context is L*H in accent 1 (fig. 3) and at the end of the compound accent 2 contour. In the curl/stød articulation of Eskilstuna, the corresponding articulation is H*L%, that is, without the initial L. I seek to document this shift as synchronic within the Eskilstuna dialect.

The speech materials investigated are taken mostly from interviews made in the 1960s. These materials are rich (83 informants), stratified for some social variables and well studied from a sociolinguistic perspective (Nordberg 1970).

Figures

Figure 1: E147, curl, H*L% accent 1
A quantitative implementation of the autosegmental-metrical model: the study of pitch and duration in Friulian

Paolo Roseano, Ana Maria Fernández Planas & Eugenio Martínez Celdrán
University of Barcelona, Laboratory of Phonetics
paolo_r@hotmail.it, anamariafernandez@ub.edu, martinezceldran@ub.edu

Since its appearance, the autosegmental-metrical model has been used to carry out many interesting researches on prosodic features. Within this framework, nevertheless, few attempts have been made to suggest a method which transforms raw numeric data in phonological structures in a way that aims at reducing the degree of uncertainty implied by the possible subjectivity of the researcher and, thus, at basing the phonological description on a more solid phonetic base.

This article aims at testing the validity of a quantitative approach to the implementation of the autosegmental-metrical model on both pitch and duration. Such approach was first developed in relation with Spanish (Martínez Celdrán et al. 2003, Fernández Planas et al. 2003). It relies on psychoacoustic thresholds, for both pitch and duration: it assumes that a difference in tone between two vowels is significant if it is higher than the psychoacoustic threshold of 1,5 semitones (Pamies Bertrán et al. 2002). A difference in duration, on the other hand, is assumed to be significant if it is higher than the psychoacoustic threshold of 1/3 of the duration of the adjacent vowel (Pamies Bertrán et al. 2006). Once determined whether a difference in either of such parameters is significant, the ToBI prosodic transcription is used to label the phonological structures.

For the purposes of this article, such approach has been tested on a corpus of 288 sentences uttered by a native speaker of Friulian, a rhaeto-romance language whose prosody has hardly been explored (Finco 2007, Roseano 2008). Such sentences have a SVO structure (the S being either a noun or a noun followed by an adjective). They are either statements or yes-or-no questions. Furthermore, they contain all possible combinations of oxytone, paroxytone and proparoxytone words.

Such sentences have been analyzed with the software Amper (López Bobo et al. 2007) in order to get the raw numeric data (in Hz and ms). Such data were then transformed into superficial and underlying phonological structures, by means of SPSS algorithms based on the above mentioned psychoacoustic thresholds. Statistical data about such phonological structures have been submitted to routine analyses and tests (frequency, chi-square, distribution of modes) with the aim of:

a) determining intonational and durational patterns,
b) pointing out causal relations between variables assumed to be independent (modality, syntagmatic position, lexical accent) and variables assumed to be dependent (phonological structures of intonation and duration),
c) test the effectiveness of the methodological approach.

These findings of the research concern both methodology of research and phonology of Friulian. As far as methodology is concerned, the quantitative approach proved to be satisfactory. In fact, the results obtained by carrying out statistical tests on phonological structures obtained by means of psychoacoustic thresholds implemented on a wide corpus are noteworthy and cohesive. As far as the prosody of Friulian is concerned, it has to be pointed out that there is a surprisingly clear relationship between modality and pitch on one side, and between vowel duration and lexical stress on the other side. Such a clear relationship can seldom be found in other romance languages (Pamies Bertrán 1993).
Figures

Relations among phonological structures, modality and lexical stress

References

Fernández Planas, A. M.; E. Martínez Celdrán (2003), El tono fundamental y la duración: Dos aspectos de la
taxonomía prosódica en dos modalidades de habla (enunciativa e interrogativa) del español, Estudios de fonética experimental, XII, pp. 165-200.


Roseano, P. (2008), El to fonamental i la durada: dos aspectes de la taxonomía prosódica en dues modalitats de
parla (enunciativa i interrogativa) del friülà, Treball d’Investigació en Tipologia Lingüística, Universitat de Barcelona.
Leading tone alignment in Occitan disapproval statements

Rafèu Sichel-Bazin
Universität Osnabrück-UPF
rsichelb@uos.de

Occitan is a Romance language that, like French, accepts pitch movements associated with unstressed syllables. Hualde (2003) notices that “the falling (or low) nuclear accent of Occitan appears to differ from that of the other languages”, arguing that very often he had found “a fall from a preceding syllable with a secondary accent”. The nuclear pitch accent of Occitan disapproval categorical statements consists in a rising-falling contour as described in Fig. 1, and the rise of this contour appears to be linked with the syllable (S-1) that immediately precedes the last stressed syllable of the utterance (S*). That’s why in this case the rise may not be an early rise or a secondary accent, but has to be considered as a leading tone configuration that is part of the pitch accent.

![Fig. 1 – Schematic representation of the nuclear pitch accent contour of Occitan disapproval statements. The white rectangles symbolize unstressed syllables (S-2 and S-1 respectively); the grey one, the last stressed syllable of the sentence (S*); and the black line, the stylized F0 curve.](image)

In this paper, we focus on the alignment of this rise with regard to the segmental string. We will proceed to the analysis in Praat of recordings of sentences obtained from a situation survey. The methodology we use, based on Prieto (2001), allows getting natural utterances in a semi-directed way by prompting the informers to express their reaction to hypothetical situations of the everyday life. The informers have to express their disapproval to 5 different situations, uttering, for each one, sentences with 12 different expressions in nuclear position: 6 paroxitonic ones and 6 oxitonic ones, all with a [CV(S-1) CV(S*) CV] final structure.

The analysis of a pilot study shows that the F0 turning points of the rise of this pitch accent are quite constantly anchored to particular segmental points of S-1: there appears to be a clear F0 turning point coinciding with the onset of S-1 which has to be interpreted as a L (low) target, and a peak occurs mainly at the end of S-1 or sometimes during the onset of S*, before pitch falls. Hence, this pitch accent has two leading tone targets L and H, both coinciding with S-1, and it ought to be labeled LH+L*.

Very little literature has taken into account the description of leading tones alignment. Pierrehumbert (1980) and Pierrehumbert and Steele (1989) affirm that unstarred tones lead or trail the starred one “by a given time interval”, and as a consequence, may spread over one or more syllables depending on the speech rate and the length of the segments. To our knowledge, only Grice (1995) proposes that leading and trailing tones may present two different systems of alignment: she claims that, whereas trailing tones occur a fixed interval of “normalized time” after the starred tone, leading tones appear to be linked to the syllable preceding the accented one (if there is one available).

It arises from our study that both L and H leading tone targets of Occitan disapproval statements are aligned with the syllable preceding the accented syllable. This represents evidence bearing out Grice’s theory about the alignment of leading tones: leading tones must not be aligned according to time but to syllables.
References

The first part of this talk is simple, the second part is experimental and builds on joint work.

We first highlight an important result of Bartels 1997 (not among the authors here) that is not fully recognized in the literature. The question, in simple terms: Why, across languages, is ‘question intonation’ [/] felt to be appropriate in yes-no-questions (Is it raining [?]) but less so in wh-questions (When was it raining [?]) and impossible in alternative questions (Is it raining or isn’t it [\]). Arguments against earlier proposals such as [?] always marking continuation (Is it raining [?] or isn’t it. Kretschmer 1938) are reviewed. The classical work by Pierrehumbert and Hirschberg 1990 does not address this issue.

The answer, building on Bartels: (a) intonational meanings operate on salient propositions (rather than on the output of compositional semantics) and (b) for a given salient proposition p, [\] commits the speaker S to p while [?] doesn’t. (Independent evidence will be shown for (a) and (b)). For all question types, a possible salient proposition is the assumption of the existence of an answer, shown in (1). S’s endorsement of the existence of an answer can be marked with [\] in all cases, and correspondingly all kinds of questions can be marked [\].

In the alternative question, this directly generates the ‘no other alternatives’ impact of alternative questions (Do you want tea or coffee[\]–i.e. you want tea or coffee) as opposed to yes/no-questions (Do you want tea or coffee[?]–i.e. perhaps not).

Importantly, now, yes/no-questions are different because there is a further salient proposition available (one not shared by the other question types): In Is it raining? this further salient proposition is It is raining. This of course must not be endorsed by S, else the questioning purpose is undermined. This salient proposition may be chosen and must then be marked by [?]. It will be shown how this account in terms of salient propositions is also compatible with variation observed in the corpus literature.

In the second part of the talk, we present the results of a combined production and perception experiment on Brazilian Portuguese intonation. Six speakers from the Campinas area read three sentences with ante-penultimate stress in five environments. Each environment is constructed to trigger one of five central contours of BP informally illustrated by Cagliari 1982 (see also de Moraes 1998): (a) statement (b) emphatic statement (c) yes/no-question (d) surprise question (e) continuation. In a following perception task, twenty different speakers from the same area matched each production to the best-suited of the five contexts. The productions that were recognized significantly well are acoustically evaluated.

Results for the well-recognized tokens: Emphatic statements and normal statements employ a H+L* L- contour and are distinguished in the phonetic height. Yes/no-questions and surprise questions are phonologically distinguished as L+H* L- vs. L*+H L-, see Fig. 1.

A decomposition of the tune meanings allows us to isolate how this intonation system simultaneously encodes two distinctions: (i) what is coded by syntactic inversion in English (declarative vs. interrogative, not syntactically encoded in BP) and (ii) additionally “[\]” (endorsement by S) vs. “[?]” from above. In BP, [?] is an L+H pitch accent and [\] is an H+L pitch accent. Declaratives are marked by a star on L, i.e. ...L*..., interrogative with the star on H, i.e. ...H*.... (This latter converges with the perception studies of de Moraes 1984 on BP.) In combination, H+L* is a declarative with [\], i.e. a statement; L+H* is an interrogative with [?], i.e. a yes/no-question. Importantly, L*+H is a declarative with [?]: The BP ‘surprise question’ turns out to be equivalent of
the English declarative question (e.g. *It is raining?* Gunlogson 2001). This equivalence will be supported by shared restrictions on their use: Like English declarative questions, BP surprise questions are felicitous only where the addressee can be expected to believe the proposition at issue (also 'confirmation question').

**Examples and figures**

\[ p = \text{existence of answer} = \]

<table>
<thead>
<tr>
<th>p</th>
<th>It is raining or it isn't raining.</th>
</tr>
</thead>
<tbody>
<tr>
<td>It was raining at some time.</td>
<td></td>
</tr>
<tr>
<td>It is raining or it isn't raining (alternative q.)</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 1.** Plots of the measurements of the tokens that were recognized particularly well in the perception experiment, in the categories yes/no-questions and surprise questions. The vertical black bars represent the extent of the sentence-final verb of the intransitive stimulus clauses. The vertical grey bars delimit the stressed syllable in the verb. The plots are normalized for F0 and for time (preserving placement of a given point in its temporal segment) and pool the relevant productions of the six speakers.

**References**


The role of pitch height in constraining the inferential space in Catalan yes-no questions

Maria del Mar Vanrell, Ignasi Mascaró, Francesc Torres-Tamarit, Pilar Prieto
U. Pompeu Fabra, Institu Menorquí d’Estudis, U. Autònoma de Barcelona, ICREA-UPF
mariadelmar.vanrell@uab.cat, ipmascaro@gmail.com, pilar.prieto@upf.edu, francesjosep.torres@uab.cat

Previous research has shown that informational meaning can be encoded intonationally in polar questions (Grice et al. 1995; Grice and Savino 1997, 2003a, 2003b; Crocco 2002; Kügler 2003; Santos and Mata 2008). This is especially interesting to note in languages where syntactic aspects are not present, and consequently, the decision of the informational-status of a certain question relies heavily upon prosodic features as the primary cues to the informational-status (Grice and Savino 2004). With regards to previous studies on this issue, Grice and Savino (1997) demonstrated that the degree of confidence with which the speaker believes the information to be shared with the interlocutor is reflected in Bari Italian in the choice of a specific pitch accent. In contrast, the results of Kügler (2003) revealed that the interaction between intonation and information structure was displayed through the use of different boundary tones.

Following Escandell’s proposal (Escandell 1993, 1998) about transactional speech (in which the speaker’s discursive intentions are mainly informative) and the real knowledge/knowledge presupposition scale (see Table 1), a production experiment was conducted in Central and Balearic Catalan in which information- and confirmation-seeking questions were elicited. The results showed that the different dialects of Catalan use basically two different strategies to encode intonationally the difference between information- and confirmation-seeking questions, that is: the use of different boundary tones (Central and Ibizan Catalan) or different pitch accents (Minorcan and Majorcan Catalan). The goal of this paper is to mirror perceptually the results obtained in production by focusing only on Majorcan Catalan. This production study showed that the main cue that Majorcan speakers used to mark the difference between information- and confirmation-seeking questions was the pitch height of the prenuclear syllable. Thus, information-seeking questions are mainly characterized by the presence of a falling nuclear accent H+L+* L% in which the leading tone H is significantly higher in pitch than in confirmation-seeking questions (Fig. 1).

In order to test whether the difference in pitch height of the prenuclear syllable is also an important cue perceptually, we undertook three types of experiments: a) a congruity test designed to evaluate whether the two nuclear configurations can be used successfully in different pragmatic contexts, b) a rating test in which subjects had to rate the degree of presupposition of different sentences on a 4 point scale and c) a standard identification task where the original stimulus was manipulated from a high to an extra high leading tone in 11 steps of 11.2 Hz each, and conversely. Pilot results with 5 listeners show: a) that listeners know very well in which context both interrogatives can or cannot be produced, b) that the real knowledge/knowledge presupposition scale is also reflected in perception (prosodic and morphosyntactic cues were successfully related to the speaker’s presupposition of the truth value of certain sentences) and c) that listeners can identify in a near-categorical way the perceived stimuli (see Fig. 2). The results of this experiment demonstrate that this prosodic feature (H in the prenuclear syllable) is the main cue encoding this type of procedural meaning. Specifically, this type of procedural meaning would constrain the space for making inferences and would let the hearer to know how to proceed by minimizing the processing effort (Sperber and Wilson 1986, 1995; Blakemore 1987; House, 2007).
Figures and Tables

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Hearer</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

Real knowledge | Knowledge presupposition

Table 1. Sketch taken from Prieto (2002) and adapted from Escandell (1993).

Figure 1. Waveform and fundamental frequency contour of information seeking-question *Tien mandarines?* (Do you have tangerines?) –left panel– and confirmation seeking-question *Has agafat sa bossa de mandarines?* (Have you taken the tangerines’ bag?) –right panel.

Figure 2. Results of the congruity test (left panel), rating test (middle panel), and standard identification task (right panel).

References


Escandell-Vidal, M.V. 1993. Introducción a la pragmática, Ánthropos, Universidad Nacional de Educación a Distancia.


Prosodic Optionality or Syntactic/Semantic Choice?
Michael Wagner
McGill University
chael@cornell.edu

Many studies have shown that prosody can disambiguate structural ambiguities, but it remains controversial how close the relation between syntax and prosody is. At least sometimes prosodic boundaries seem to be optional. Some studies found that speakers only employ prosody if they are aware of the ambiguity (Snedeker & Trueswell 2003), but others found that prosody reflects attachment even if not (Schafer et al. 2000, Kraljic & Brennan 2005). Based on the prosody of arithmetic formulas, this talk asks the question whether distinct options in prosody do not in fact reflect choices between different structures that differ in their syntax and semantics. Arithmetic formulas, e.g., are often prosodically disambiguated by boundaries of different strength (e.g., O'Malley et al. 1973, Streeter 1978):

(1)  a. (A + B) * C   ‘left-branching’
    b. A + (B * C)   ‘right-branching’

However, they can also be produced with a ‘flat’ prosody. If prosodic disambiguation is optional, these cases can be interpreted as renditions of (1a,b), in which the speaker chose not to disambiguate. The alternative hypothesis pursued here is that flat renditions in fact reflect a third structure, which involves a list of calculation instructions, typically set apart by boundaries of equal strength:

(2)  A, | + B, | * C   ‘flat’

A list is intuitively interpreted from left to right, in this case this leads to a result equivalent to left-branching. Flat prosody should not be compatible with a right-branching formula, so prosodic optionality makes different predictions.

We tested this in 5 experiments. First, we synthesized 16 formulas, adding pre-boundary lengthening and pause duration at either the first or the second or both boundaries. In a forced choice experiment (Exp 1) listeners could indeed distinguish 3 structures, but perform at chance when there are no boundaries. In Exp. 2 and 3 subjects they had to calculate the result of the formula. The results show that the flat prosody most compatible with a list structure (1_1) leads as predicted to the same result as the ‘left-branching prosody’ (0_1). Prosody competes with other sources of disambiguation such as order of operation and a strong left-to-right bias if subjects are not explicitly instructed that prosody disambiguates the bracketing.

Two production experiments further test the hypothesis. In experiment 4, it was obvious that there are two different bracketings that should be disambiguated, in experiment 5, subjects only ever encountered one bracketing, either left- or right-branching, and the ambiguity was less obvious. In Exp. 4, all 16 participants generally disambiguated. In Exp. 5, 2 out of 4 participants run so far on only left-branching formulas used predominantly flat prosody, but all 4 recorded only on right-branching structures used a right-branching prosody. Both perception and production results suggest that a flat prosody reveals choosing a list structure over a nested structure, rather than giving evidence for prosodic optionality.

Note: 16 subjects participated in each Experiment.
1 = Left-Branching; -1 = Right Branching; 0 = Flat
Figures

Experiment 1: Forced Choice between three bracketings

Experiment 2 & 3: Participants Calculate the Result of the Formula

Experiment 2: No Instruction, strong left-to-right bias:

Experiment 3: Instruction that the strong formulas are pronounced such that they reveal the bracketing, less bias:

Order of Operation Effect in Experiment 2, but not Experiment 3:

First operation plus/minus: 3 + 4 * 5
First operation times: 3 * 4 + 5
Intonation in Discourse: Gradient or Categorical Behavior?

Margaret Zellers & Brechtje Post
Research Centre for English & Applied Linguistics, University of Cambridge
mkz21@cam.ac.uk, bmbp2@cam.ac.uk

A main problem in prosody research is the attempt to separate out the linguistic from the paralinguistic aspects of intonation. Ladd (1996:36) has argued that gradience is the “defining characteristic” of paralanguage, whereas linguistic structures are categorical. However, especially within the realm of intonational phonology, categorical structures have proved problematic to identify. In particular, the attempt to isolate such structures within the context of discourse above the level of the intonational phrase (IP) has thus far received relatively little attention in intonational phonology.

One area in which we can try to observe the behavior of intonation at a level higher than the IP is with reference to the topic structure of the larger discourse. In theories of topic structure, we encounter a similar problem to that which intonation faces in trying to identify categorical versus gradient behavior. Some theories that have been proposed to account for the topic organization of discourse would suggest a gradient interpretation of topicality, while others involve a more categorical approach. An example of a topic structure theory that takes a more gradient view is the hierarchical analysis proposed by Grosz & Sidner (1986). Since the simple hierarchy they propose is potentially infinitely recursive, this theory lends itself to a gradient pattern of variation, dependent upon the amount of active information. An alternative model is provided by Nakajima & Allen (1993) and Wichmann (2000), who propose that topic structure can be described by four discontinuous categories related to the relationships between information contained within individual units (see Fig. 1 for examples of categories). It must be pointed out that it is not yet entirely clear what these categories might consist of: Wichmann (2000) proposes that the categories are to do with the relative newness of the information in the unit, while Brazil (1997) suggests that topicality might be more accurately described in terms of the relative independence of a unit from surrounding units.

Although both of these approaches to the organization of discourse topic have merit, thus far there has been very little systematic evidence provided to address the issue of whether gradient or categorical behavior is evident in this domain. A production study by Zellers (submitted) addressed this lack of data from an intonational point of view. She investigated the variation of the span of H*L (falling) pitch accents in the first position in an IP as produced by speakers of Standard Southern British English (SSBE). The pattern apparent in this cue (see Fig. 1) was more in line with a categorical theory of topic structure than a hierarchical theory, and was furthermore independent of the shifting of the overall pitch range with regard to the position in the utterance group (supra-declination) which falls demonstrably outside of the linguistic system.

The implication of this result is that topic structure as evidenced in the prosodic organization of spoken language is potentially categorical in nature, and therefore not paralinguistic but rather part of the linguistic system. Furthermore, at this point it is difficult to identify where these categories reside. On the one hand, the differences in span could result from an event or feature specification in the intonational phonology itself; alternatively, it could be the direct realization of a different level of linguistic structure, specifically the topic structure of the discourse. The free variation in the order of the categories in the discourse suggests that these categories are discoursal, rather than intonational; but more evidence is required to confirm or deny this tentative conclusion. In either case, intonational theory will need to develop an account of the influence of an additional layer of categorical structure beyond the level of the IP.
Figures

![F0 Fall Range - Schematic](image)

Fig. 1 Span of F0 falls (schematic) from Zellers (submitted).

References


