

UNIVERSIDADE ESTADUAL DE CAMPINAS  
INSTITUTO DE ESTUDOS DA LINGUAGEM

# Cadernos de ESTUDOS LINGÜÍSTICOS

43

Organizado por:

*Plínio Almeida Barbosa*

**ANÁLISE E MODELAMENTO  
PROSÓDICOS DA FALA**

Cad.Est.Ling.	Campinas	Nº 43	p.1-150	Jul./Dez.2002
---------------	----------	-------	---------	---------------

## **UNIVERSIDADE ESTADUAL DE CAMPINAS**

Reitor: Carlos Henrique de Brito Cruz

Vice-Reitor: José Tadeu Jorge

## **INSTITUTO DE ESTUDOS DA LINGUAGEM**

Diretora: Charlotte Marie Chambelland Galves

Diretora-Associada: Márcia Azevedo de Abreu

## **SETOR DE PUBLICAÇÕES**

Coordenadora: Maria Eugênia Boaventura

## **EQUIPE EDITORIAL**

J.A. Duek/L.A. Santos/E.A. Santos

Capa-Projeto: J.A. Duek

Layout e Arte Final: E.A. Santos/L.A. Santos

**CADERNOS DE ESTUDOS LINGÜÍSTICOS** é uma publicação semestral do Departamento de Lingüística do Instituto de Estudos da Linguagem da UNICAMP. A revista aceita colaboração de pesquisadores de outras Instituições, publicando estudos em português, espanhol, inglês ou francês. Os trabalhos, acompanhados de resumos, serão submetidos ao Conselho Editorial.

Para remessa de originais, aquisição de números avulsos e assinaturas, dirigir-se a

**UNICAMP/IEL**

**Setor de Publicações**

**Caixa Postal 6045**

**13084-971 - CAMPINAS - SP - BRASIL**

**Fone/Fax.: (19) 3788 1528**

**e-mail: *spublic@iel.unicamp.br***

**<http://www.unicamp.br/iel/>**

**PEDE-SE PERMUTA**

**SUMÁRIO**

Foreword .....	5
<b>OSAMU FUJIMURA</b>	
Temporal organization of speech utterance: a C/D model perspective .....	9
<b>GERÁRD BAILLY &amp; BLEIKE HOLM</b>	
Learning the hidden structure of speech: from communicative functions to prosody .....	37
<b>FRED CUMMINS</b>	
Entraining speech with speech and metronomes .....	55
<b>PLÍNIO A. BARBOSA</b>	
Explaining Brazilian Portuguese resistance to stress shift with a coupled-oscillator model of speech rhythm production .....	71
<b>MICHAEL O'DELL &amp; TOMMI NIEMINEN</b>	
How long is a stress group? .....	93
<b>SANDRA MADUREIRA</b>	
An acoustic study of phonological phrases containing sequences of words with adjacent primary-stressed syllables: does stress shift occur in Brazilian Portuguese? .....	109
<b>AGLAEL J.A. GAMA-ROSSI</b>	
Considerations on some aspects of the relationship between intrinsic and extrinsic time in two 4-year-old-children's and one adult's speech for duration in Brazilian Portuguese .....	127
Publicações recentes, recebidas através de permutas .....	143

## FOREWORD

### TOWARD DYNAMICAL, SYLLABLE-RELATED MODELS OF SPEECH PROSODY: DATA ON FIVE LANGUAGES

The idea of organizing this collection of seven papers was twofold. First, to bring together, in a single volume, researchers who not only have devoted their work to analyzing prosodic form and content according to rigorous experimental protocols, and further more crucially, to bring together researchers who assume a modeling approach by proposing phonetic or phonological models for utterance implementation. Second, to present evidence for universal and language-specific prosodic properties by gathering information on common and specific prosodic properties from five languages, not all genetic-related: (American and Irish) English, Japanese, French, Finnish, and Brazilian Portuguese.

All papers share a dynamical, linguistic view of prosody according to which the phonetic component does not realize a blind, straightforward language-independent mapping from (language-specific) grammar to sound, but it is highly language-dependent as well as highly constrained by bio-mechanical properties of the speech production system. Some also share a more radical view, similar to Browman and Goldstein's (papers by Barbosa, Madureira, and Gama-Rossi), that of a gestural, dynamical phonological specification.

Perhaps the most striking shared assumptions in this issue are couched on two underlying hypotheses. The first considers that the building blocks of utterances' organization are syllable-sized units acting as beats, most likely anchored at the vicinity of the vowel onset (or the perceptual-center. Cf. Morton, Marcus & Frankish, 1976; Pompino-Marschall, 1989; Janker, 1995). This can easily be grasped by rapid inspection of Fujimura's notion of base function in the paper that opens this volume. The second hypothesis considers a multileveled approach of prosody specification: base function and consonantal elemental gestures with the separation of metrical and tonal features in Fujimura (cf. Öhman, 1966); superposing levels in Bailly and Holm; phrase stress/metrical and syllabic domains in Cummins, Barbosa, O'Dell and Tieminen, Madureira and Gama-Rossi.

*Temporal Organization of Speech Utterance: A C/D model perspective*, the last paper before Osamu Fujimura's retirement, was written especially for this issue, which is a great privilege. This work, with boldly clear contents, discusses the well-known C/D model as a language-specific framework of phonetic implementation, having, as a skeleton, a syllable and boundary train of pulses (base function), supplemented with

consonantal elemental gestures (tonal features can be included and completely specified in the base function as a variable independent from stress). By separating a component based on a vowel continuum from the around-syllable-margin consonantal gestures, which introduces local perturbations onto the vocalic continuum, his model is able to explain the utterances' stress and pitch-related patterns. His view presupposes a linguistic phonetics separated from classical phonology. The former is able to handle most part of the allophonic variation, receiving as input an abstract featural specification from a phonological inventory. His proposal is exemplified with data from two unrelated languages: American English and Tokyo Japanese.

*Learning the hidden structure of speech: from communicative functions to prosody*, by Gérard Bailly and B. Holm, reminds us the breakthrough represented by Elman & Zipser's seminal, homonymous paper (1988). The French breakthrough is couched on the ICP morphogenetic model (for "generation of global shapes") of utterances' prosodic implementation, which is improved here. Their model allows to automatically generate intonational forms according to different communicative functions. This is done in two steps. First, by decomposing the raw melodic shapes into prototypical contours associated to a small set of predefined domains, which are used to train a recurrent neural network. Second, by generating the utterance's contour from a rich linguistic description of the utterance using the same neural network. After generation, the prototypical contours are further combined to produce the intended contour for the utterance. A notion of the crucial aspects of global melodic contours in explaining subjects' performance is implicit in their work and is related to the expectancies that subjects have about intonational structure. The model was developed for French but its extension to Galician is on the way (maybe another domain for the prototypical contours could emerge).

In *Entraining Speech with Speech and Metronomes*, Fred Cummins presents two novel approaches to studying speech: the Speech Cycling Task and Synchronous Speech. The first technique is known from the work by Robert Port, van Gasser and Fred Cummins himself, who presented a first incursion into Temporal Phonology, by proposing as a research program to explain rhythmic cross-linguistic differences from differences in underlying dynamical, language-specific rhythmic systems. This technique consists in asking subjects to try to adjust the phase of the occurrence of a stressed syllable relative to the beginning of a repeated utterance when listening to specific external stimulation. This technique traces back to Stetson (1951 [1928]). The second technique is entirely new and consists in asking subjects to read texts simultaneously, face-to-face or not. Both techniques allow to investigate the inner nature of speech rhythm, which emerges to be highly dynamical. Cummins worked with American and Irish English.

In *Explaining Brazilian Portuguese resistance to stress shift with a coupled-oscillator model of speech rhythm production*, I present and implement a dynamical model of speech production, based on coupled-oscillator theory. This model is meant to capture the variability of Brazilian Portuguese behavioral data on acoustic duration by posing two levels of timing: an extrinsic, properly rhythmic level implemented by the coupling of a phrase stress oscillator (metrical level) and a syllabic oscillator (a

pacemaker), and a intrinsic level specified in a gestural lexicon. The model receives as input syntactic-semantic and lexical descriptors. It is capable of simulating with great precision (compared to natural data) the consequences for duration of a complex phonological phenomenon such as stress shift. It is argued that languages using duration to signal phrase stress would not exhibit phenomena such as stress shift without the use of other acoustic parameters. It is also argued that, even when exhibiting stress shift, duration is controlled exactly in the way predicted by the model.

In *How Long is a Stress Group?* Michael O'Dell and Tommi Tieminen explore a suggestion I made in a previous paper, that one stating that the number of syllables in a stress group can be influenced by the requirements of keeping phrase stress and syllable oscillators synchronized. They show that the length of a stress group is not only dependent on external control but is a function of the dynamical forces operating between the two coupled oscillators. If they are true, it has at least an important consequence: the interplay between seemingly highly "symbolic" linguistic levels and stressing and syllabicity is bidirectional, which means that the description of higher linguistic levels should be commensurable to the continuous variables operating on the rhythmic level. The language studied is Finnish.

In *An Acoustic Study of Phonological Phrases Containing Sequences of Words with Adjacent Primary-Stressed Syllables: does stress shift occur in Brazilian Portuguese?* Sandra Madureira uses three Brazilian Portuguese corpora to investigate the control of pitch in clash and non-clash conditions. Her results show clearly that, in general,  $f_0$  evolution in both conditions is quite similar and confirm my findings by discussing pitch patterns in relation to the organization of the entire utterance and to a new way of looking at the notion of stress timing. Her findings point to a way of integrating intonation and rhythm into a general model of prosody production.

In *Considerations on some Aspects of the Relationship between Intrinsic and Extrinsic Time in Two 4-Year-Old-Children's and One Adult's Speech for Duration in Brazilian Portuguese*, Aglael Gama Rossi analyses data on Brazilian Portuguese acoustic duration in four-year old children and adults under the light of Barbosa's dynamical model of rhythm. The picture that emerges clearly from her data is that the children are not contrasting reduced and full syllable-sized units as adults do, and also that they seem to use a much poorer hierarchy of prosodic constituency. In order to account for these data, she proposes that children have a more regular phrase-stress oscillator than adults (and possibly partially – or weakly - coupled to a syllabic oscillator). Her data also support the idea of the subsequent acquisition of a language-specific coupling strength between a still highly uncoupled syllabic oscillator and a phrase stress oscillator more intimately related to higher grammatical levels.

Taken together, the results of the analyses and models presented here do sustain a dynamical view of prosody production for which the language-specificity and the linguistic control of fine-grained patterns of duration and  $f_0$  *gestalten* emerge very clearly. It is now up to the readers to judge the volume from the pieces of evidence presented therein.

Before closing I would like to acknowledge my debt to Eleonora C. Albano, who encouraged me to organize this issue. I also vividly thank all authors for their kind

participation in this collective work and to wish that our interaction from now on be more dynamical than ever.

I translated the original abstracts at the top of the six first papers to Portuguese. I assume the risk for any inconsistency resulted from this translation.

---

## REFERENCES

- ELMAN, J.L. and ZIPSER, D. (1988). Learning the hidden structure of speech. *J. Acoust. Soc. Am.* 83, 1615-1626.
- JANKER, P.M. (1995). On the influence of the internal structure of a syllable on the p-center perception. *Proceedings of the XIII International Congress of Phonetic Sciences*. August 13-19, Stockholm, Sweden, 2, 510-513.
- MORTON, J. MARCUS, S. and FRANKISH, C. (1976). Perceptual-centers (P-centers). *Psychological Revue.* 83 (5), 405-408.
- ÖHMAN, S. (1966). Coarticulation in VCV utterances. *J. Acoust. Soc. Am.* 39, 151-168.
- POMPINO-MARSCHALL, B. (1989) On the psychoacoustic nature of the P-center phenomenon. *Journal of Phonetics.* 17, 175-192.
- STETSON, R. H. (1951). *Motor phonetics: a study of speech movements in action*. 2nd ed. Amsterdam: North-Holland.

Plínio A. Barbosa  
Campinas, August 15<sup>th</sup> 2002