SÃO PAULO SCHOOL OF ADVANCED STUDIES IN SPEECH DYNAMICS

FAPESP

June 7-11, 2010

ABSTRACTS

Prefácio Design - www.prefacio.com.br

dinafon

UNICAMP

Sponsors:

FAPESP

CNPq

São Paulo School of Advanced Studies in Speech Dynamics
SPSASSD 2010
Accepted Papers

Organizing committee

Eleonora Albano *Unicamp*

Didier Demolin *USP/ULB*

Beatriz Raposo de Medeiros *USP*

Rui Rothe-Neves *UFMG*

Administrative committee

Marilda Gifalli *USP*

Cláudia R. Tavares *USP*

Antonio Pessoti *Unicamp*

Denize Pozzani *Unicamp*

Scientific committee

Alice Turk *University of Edinburgh*

César Ades *USP*

Didier Demolin *USP/ULB*

Eleonora Albano *Unicamp*

Ioana Chitoran *Dartmouth College*

Hani Camille Yehia *UFMG*

John Kingston *University of Massachusetts*

Jorge Lucero *Universidade de Brasília*

Khalil Iskarous *Haskins Laboratories*

Osvaldo de Oliveira Jr. *USP–São Carlos*

Thaïs Cristófaro-Silva *UFMG*

Scientific committee - Revisors

Aglael Gama Rossi

Alice Turk

Bernard Harmegnies

César Reis

Claire Pillot Loiseau

Didier Demolin

Eleonora Albano

François Pellegrino

Hani Camille Yehia

Hans Van de Velde

Ioana Chitoran

Ivone Panhoca

Jacqueline Vaissière

John Kingston

Jorge Lucero

Khalil Iskarous

Natalie Operstein

Nathalie Vallée

Osvaldo de Oliveira Jr.

Pascal Perrier

Regina Célia Fernandes Cruz

Ricardo Molina de Figueiredo

Solange Rosato

Stéphanie Weisser

Thaïs Cristófaro Silva

Sumário

A word from the organizing committee

When a final dot is put in this document, there will be less than a year that the organization of the São Paulo School of Advanced Studies in Speech Dynamics has started. As soon as Eleonora Albano, Didier Demolin and the members of the Dinafon research group took cognizance of the *FAPESP São Paulo School* call for proposals, most researchers and institutions involved in the study of speech in the country joined forces to make this event possible. The speed at which the project, its evaluation, the call for contributions, the invitation of foreign colleagues and the organization of the event have taken place is unprecedented.

This school is a great opportunity to show that Brazil has a very lively community of students, researchers and faculty members involved in various aspects of the study of speech dynamics. Brazilian students will have a unique opportunity to meet both faculty and students from abroad. The event will also foster the creation of networks among participants, and open countless new opportunities to everyone involved. For both the participants and their institutions, there could be no better promise of continuity.

We are also very happy to thank and welcome the colleagues who accepted our invitation to teach and participate in the event. Their contribution is fundamental for the SPSASSD goals to be reached. Their expertise and enthusiasm helped us meet very high standards in the program.

We are also very happy to welcome the students who are coming from Europe and South and North America. We hope that the school will enhance their training and help them discover the academic potential of Brazil.

The members of the organizing committee are particularly happy to thank FAPESP for giving us the opportunity to make this event come true. The important investment made by FAPESP will definitely change the landscape of speech studies in the country. It will also help the Brazilian community of speech studies to dialogue with colleagues from various parts of the world. We also wish to thank all the other institutions hosting or supporting the project: IEA-USP (Instituto de Estudos Avançados da Universidade de São Paulo), CNPq, and IEL-UNICAMP (Instituto de Estudos da Linguagem da Universidade Estadual de Campinas).

Many people and colleagues have made this event possible. We are particularly thankful to the members of the various committees, especially those of the scientific committee, who made fast, effective reviews, sometimes in a very short time. We finally wish to express our gratitude for the daily dedication of Marilda Gifalli, Claudia R. Tavares, Antonio Pessotti, and Denise Pozzani.

Welcome all! And do have a good time in São Paulo!

Eleonora Albano, Didier Demolin, Beatriz Raposo de Medeiros ,Rui Rothe-Neves.

SPSASSD 2010 Accepted Papers

Aline OLIVEIRA. **Fundamental frequency behavior in sad and happy speech**

Amelia KELLY. **Join cost optimisation in unit selection speech synthesis**

Ana Paula P F ENGELBERT . **Differences in the production of english words of sequences cvc and cvcv by brazilian learners: an acoustic analysis**

Ana Paula ROZA. **The relationship between speech and singing in Bossa Nova**

Anne ZIMMER-STAHL. **Speech Timing, Fluency, and Tonal Alignment**

Antonio PESSOTTI. **Similarities and differences between sung and spoken speech.**

Argyro KATSIKA. **The effect of prominence position on boundaries in Greek**

Benjamin PARRELL. **How /b, d, g/ differ from /p, t, k/ in Spanish: A dynamic account**

Celine DE LOOZE. **ADoReVA and ADoTeVA, Two PRAAT Plugins for the Automatic Detection of Register and Tempo Variations**

Christin KIRCHHUEBEL. **The effects of Lombard speech on vowel formant measurements**

Claire NANCE. **Aspiration in Scottish Gaelic stop consonants**

Clerton Luiz Felix BARBOZA. **An Introductory Analysis of the Roles of Duration, Intensity, and F0 in Explicit and Implicit Performatives in Brazilian Portuguese**

Darja APPELGANZ . **Word-Final Devoicing in Russian and German: Elctroglottographic Analysis**

Denise Pozzani de F. BARBOSA. **Gradient allophony in Brazilian Portuguese alveolar stops: data from a dialect spoken in Jundiaí-SP**

Diego JIQUILIN-RAMIREZ. **Approaches to the synchrony and diachrony of back fricatives in three dialects of Spanish**

Fabiana PENIDO , Renata NORONHA, Kisy CAETANO, Marisa MARISA, Camila DI NINNO and Ana Teresa BRITTO. **Correlation between the findings on the nasal air emission test and nasopharyngoscopy in patients with operated cleft lip and palate**

Fernando CARVALHO. **Prosodically-induced variation in acoustic parameters of oral vowels in Tikúna**

Francisco MENESES and Vera PACHECO. **Instrumental analysis of vowel and plosive duration in brazilian portuguese**

Gustavo Augusto Mendonça ALMEIDA. **Segmental reduction in sequences of high vowels**

Jaime CRISOSTO. **About the linguistic consciousness of prestige and stigmatization of /tʃ/**

allophones in first grade children belonging to schools of different socioeconomic groups in Concepción, Chile.

Jeremy DAY-O'CONNELL. **At the intersection of music and language: stylized intonation**

Jill THORSON. **Gestural and acoustic evidence for the syllabic organization of children's early words**

Larissa RINALDI. **Procedures for the Analyses of Child Vowels and Obstruents in Brazilian Portuguese**

Laudino ROCES. **Vowel-to-vowel perseveratory dissimilation effect in Brazilian Portuguese mid vowels**

Leandro BOLANOS. **Aspiration of /s/ in Spanish. An account of morpheme boundary processes.**

Leonardo LANCIA and Susanne FUCHS. **Applying Recurrence Analysis to Speech**

Lisa ROBERTS. **Authentic and acted responses to physical and emotional distress**

Luci KIKUCHI . **High vowels in Brazilian Portuguese and British English**

Luciana Lessa RODRIGUES. **Phonological relations between speech and writing in pre-school children**

Luciana PAGAN-NEVES and Haydée WERTZNER. **Elettoglotographic study from Brazilian Portuguese fricative voiced sounds**

Lucila Rey Rocha SCHLIEMANN. **Analysis of voiced and voiceless consonant production in children between 6-8 years old with and without speech disorders**

Ludimila NASCIMENTO, Stela LEMOS and Rui ROTHE-NEVES. **The influence of environmental noise on the performance of students in tests of tonal pattern of frequency and duration**

Magnun MADRUGA. **$C_{nasal}V$ and $C_{stop}V$ biases in Brazilian Portuguese acquisition: an exploratory study**

Marcia R. BECKER. **Production of [m] and [n] in Codas by Brazilian Students**

Márcio PINTO . **Nasality as an acoustic parameter for speaker identification**

Marco FONSECA and Mário SILVA. **The Emergence of Some Phonological Patterns in the Verbal Morphology in Brazilian Portuguese**

Marcus Vinicius Moreira MARTINS. **Frequencies Scales: an intonation and perception's speech approach.**

Maria CANTONI. **An experimental study on the perception of Brazilian Portuguese lexical stress**

Marianna NADEU. **Secondary Stress in Catalan**

Marta Helena BANDEIRA. **The dynamics of interlinguistic transfer of VOT patterns in multilingual children**

Michael COLLEY . **Near merger or lexical split? The case of mas and mais in Brazilian Portuguese**

Michael PROCTOR. **Towards a Dynamic Computational Model of Portuguese Liquid Consonants**

Priscila CÔRTES. **Factors of disagreement in the speech segmentation of the C-ORAL-BRASIL corpus**

Renne ALEGRIA and Maria Ines NOGUEIRA. **Speech Production of mild Alzheimer's disease patients**

Ricardo F. NAPOLEÃO. **Emergent Consonantal Clusters**

Rosane GARCIA and Márcia ZIMMER. **CV phonotactical biases in Brazilian Portuguese acquisition**

Sabrina GEWEHR-BORELLA, Márcia ZIMMER and Ubiratã ALVES. **Comparing the 'writing production', 'perception' and 'oral production' tripod in monolingual and bilingual data**

Scott SADOWSKY . **Phonetic characterization and socio-linguistic stratification of the vowel allophones of the Spanish of Concepción (Chile)**

Tareq MAITEQ . **Prosodic constituent structure and the spread of anticipatory pharyngealisation in two Arabic dialects**

Victor Hugo M. SOARES and Liliane P. BARBOSA. **On the nature of raised mid front vowels in Brazilian Portuguese**

Vikram RAMANARAYANAN. **Prosodic Variation within Speech Planning and Execution**

Viviana Vergara FERNÁNDEZ. **The production of [v] as an allophone of /b/ in pre-literate children in the province of Concepción (Chile)**

Fundamental frequency behavior in sad and happy speech

Aline Oliveira
alinemarafono@gmail.com
FFLCH – USP

This study aims to understand pitch behavior on emotional speech, particularly on *happy* and on *sad* speech in Brazilian Portuguese and it is part of a dissertation in progress.

Theoretical bases of this dissertation that includes four emotions (happiness, sadness, fear and anger) were found in Banse and Scherer (1996) and Erickson et al. (2006).

Subjects who participated in the study were three Brazilian actresses, with an average professional experience of 20 years. They read a text (consisting of 126 words), interpreting *happy* and *sad* in speech. Each actress repeated the text five times for each emotion. The corpus was formed by three sentences, chosen as the initial, the middle and the final sentence in the text:

Sentence 1: As células do sangue que fabricam anticorpos, são individualizadas.*(The blood cells that produce antibodies, are individualized).*

Sentence 2: As células do fígado são provavelmente iguais entre si. (*Liver cells are probably identical to each other).*

Sentence 3: Este conjunto constitui um clone linfocitário. *(This set is a clone of lymphocytes).*

Each of the three sentences was reduced to smaller units (syllables), and for each unit, F_0 was identified. This same procedure was performed for all repetitions. Free software Praat 5.1.23 and his related scripts obtained F_0 measures.

For the present work, which is part of a larger one, descriptive statistics in the Excel 2007 was run. However, in the future, statistical analysis will be run in R program and its applications.

Results: Figure 1 shows that the happy pitch contour is higher than sad pitch contour. (redo this phrase)

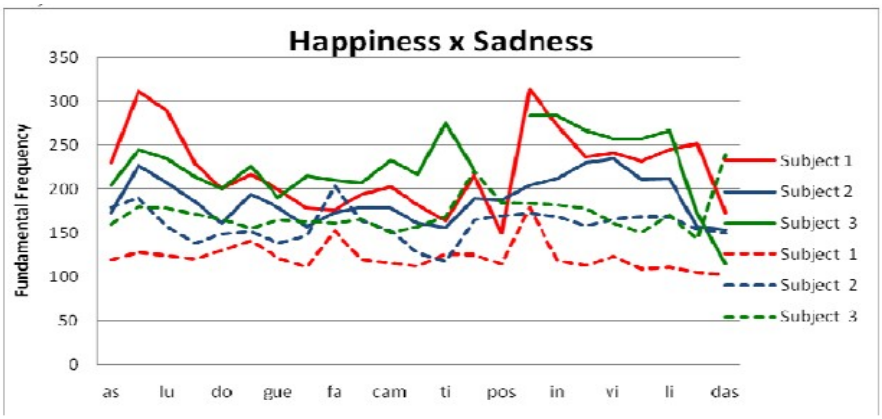


Figure 1: Each line corresponds to the average of five repetitions of the first sentence. The continuous line refers to the happiness and the sadness to the dotted line. In happy speech, Subject 1 showed a pitch range between 402 Hz and 109 Hz, and a low pitch range (254 Hz to 80 Hz) in sad speech. The same behavior was observed with the other subjects, as we can see in Table 1:

	Subject 1	**Subject 2**	**Subject 3**
Average	214	237	246
Standard deviation	46	58	63
Maximum	307	422	451
Minimum	125	139	159
Coefficient of variation	22	24	25

Table 2: Descriptive statistical data of sentence 2 in *happy* speech. Values in Hertz.

	Subject 1	**Subject 2**	**Subject 3**
Average	137	172	182
Standard deviation	49	24	32
Maximum	239	211	248
Minimum	90	78	91
Coefficient of variation	36	14	18

Table 3: Descriptive statistical data of sentence 2 in *sad* speech. Values in Hertz.

	Happiness	Sadness	S
Subject 1	222	123	p=0
Subject 2	188	159	p=0
Subject 3	227	172	p= 0

Table 4: F_0 mean values (Hertz) for each emotion per subject. S indicates significance.

Table 4 shows a p-value equal to zero (alfa = 0.05), as the result of a t-test, what indicates difference between happiness F_0 and sadness F_0. Thus, F_0 showed to be a good parameter to distinguish the emotions studied up to present. Sadness had a tendency to lower fundamental frequency and the opposite happened with *happy* speech. Considering amplitudes of mean F_0 for both emotions and looking for pitch contours in Figure 1, it is possible to assume that emotional speech can be also distinguished by its contour. However we also assume that the contour issue deserves better understanding as we intend to conclude its behavior in future studies.

Bibliography

BANSE, R.; SCHERER, K. R. Acoustic profiles in vocal emotion expression. Journal of Personality and Social Psychology. 70 (3), 614-636, 1996.

ERICKSON, D.; YOSHIDA, K.; MENEZES, C.; FUJINO, A.; MOCHIDA, T.; HIBUYA, Y. Exploratory Study of Some Acoustic and Articulatory Characteristics of Sad Speech. Phonetica, Freiburg: Karger, 2006.

Join cost optimisation in unit selection speech synthesis

Amelia Kelly
kellya16@tcd.ie

This paper outlines the research being conducted at graduate level in the area of unit selection speech synthesis. The aim of this project is to improve unit selection speech synthesis by investigating methods of optimising the join cost calculation. The join cost is a measurement of the perceptual quality of the concatenation of two consecutive units in a synthesised sentence. The higher the join cost, the less likely that combination of units will be chosen as the final synthesised utterance and so the join cost calculation has a direct effect on the quality of the synthesis. The join cost consists of a number of weighted sub-costs, normally the amplitude difference, f0 difference and spectral distance between frames of speech on either side of the join. This study aims to improve the join cost calculation first by ensuring that for each phonetic category the method of measuring spectral distance correlates well with human perception of the discontinuity, and finally by adjusting the weights associated with the sub-costs for each phonetic category. The phonetic categories examined in this study are vowels diphthongs, glides, nasals, stops, fricatives and affricates.

The first part of the study involves designing a listening test so as to determine whether joins are perceptible by human listeners, and then using different spectral distance measures on the same joins to see which automatic objective measure correlates best with the perceptual test results. This provides an automatic method of detecting spectral discontinuities and also a means evaluating how perceptually similar two sounds are. The second part describes an attempt to adjust the weights associated with the concatenation sub-costs using a genetic algorithm, with the objective measurement of spectral discontinuity as a fitness function for each phonetic category.

The experimental procedure adopted in the first part of this study closely follows that of Stylianou and Syrdal (2001). For each phonetic category, a set of six rhyming test words is created in which only the beginning of the words differ and the vowel nucleus and end phoneme are the same (eg. wave, save, rave, cave, nave, pave to examine join that occur in the middle of fricatives, in this case, the /v/ phoneme). These words are recorded in the carrier sentence “Say _ again” by a male native speaker of English. The sentences are then split into two halves at a point in the middle of the phoneme belonging the the phonetic category under examination. For example the word

“wave” belongs to the category “fricatives” and so the sentence “Say wave again” is split at the mid-point of the /v/ phoneme (as determined by the minimum of the cross-correlation function). For each group of 6 rhyming words, the first part of each utterance “Say _ again” is combined with the end part of the rest of the utterances in the group, making a total of 5 fabricated versions of the utterance and one original version for each word. The pilot perceptual test contains one rhyming group per phonetic category, and 252 stimuli overall, but a larger number of rhyming groups for each phonetic category is currently being considered. The listeners decide whether a join is present in these test stimuli and if so, if it is very, moderately or barely noticeable. Various acoustic measurements are then taken on either side of the join point and spectral distance measures are used to compute the distance between them. Examples of spectral distance measures investigated are the Euclidean distance between Melfrequency cepstral coefficients and the Kullback-Leibler distance between power spectra. For each category the spectral distance measure that correlates best with the results of the perceptual test is chosen to calculate join cost during synthesis and for use in the second part of the experiment, the genetic algorithm.

The genetic algorithm is an automatic optimisation technique used in evolutionary computing and has been shown by Alías and Llorà (2003) to be a good determiner of weight configurations for the unit selection cost function. This study proposes the use of a genetic algorithm to determine weight configurations for concatenation sub-costs for each of the phonetic categories of speech sounds previously mentioned. The genetic algorithm will make use of the spectral distance measures computed for these phonetic categories to evaluate the fitness of weight configurations. The genetic algorithm begins with a *population*, which is a table of *individuals* or vectors containing the weights of each sub-cost (eg. V={0.5, 0.25, 0.25} in this case corresponding to spectral distance, f0 distance and amplitude distance). This population table represents the first generation of possible weight configurations. For each individual (set of weights), a test sentence (containing one join, eg. /@-v v-@/) is synthesised using that weight configuration. The synthesised utterance that results is compared to the natural speech (held out of corpus) and the distance is calculated between them using the acoustic measure that has been decided upon for that phonetic category (eg. fricatives). The smaller the distance, the closer the synthesised utterance is acoustically and perceptually to the original. Therefore each individual in the population can now be represented by a number, which represents its *fitness*. The next step is *crossover*. Two individuals are selected from the population and some of their values are swapped so as to give two new vectors that are combinations of the originals. The two individuals are selected randomly from the population. However, the individuals with a higher fitness are given more weight in the population and are therefore more likely to be chosen in a random selection and to reproduce more often than

their less fit counterparts. The new vectors that result form the crossover operation are called *offspring* and they make up the first two individuals of the new population, the second generation of the genetic algorithm. Crossover continues until the second generation is complete and then the process begins again. This continues for a set number of iterations or until it converges. A different configuration of sub-cost weights can be calculated for each phonetic category.

The final step in this project is to implement these modifications in a unit selection speech synthesis system, where a different join cost spectral distance measure and different weight configurations are used depending on the phonetic category of the phoneme in which the join occurs. This new system should then be evaluated in comparison to current systems.

References

[1] Stylianou, Y. and Syrdal, A. K., "Perceptual and objective detection of discontinuities in concatenative
speech synthesis", Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, Salt Lake City, USA, 2001.

[2] Alías, F., and Llorà, X.,"Evolutionary weight tuning based on diphone pairs for unit selection speech synthesis", Proceedings of Eurospeech, 2003.

Differences in the production of english words of sequences cvc and cvcv by brazilian learners: an acoustic analysis

Ana Paula P F Engelbert
anapaulapetriu@gmail.com

This research presents an acoustic analysis of CVC and CVCV English words by Brazilian learners. The main objective in this analysis was to characterize the differences between words ending in a single consonant – an oral stop – and the same words added the grapheme –y, as in pop and poppy. According to Baptista & Silva Filho (1997) and Koerich (2002), Brazilian learners may insert an epenthetic vowel after oral stops in coda position due to the influence of Brazilian Portuguese (BP) phonotatics. If epenthesis is a common phonological process in the Brazilian learners' production, how do they differentiate English words ending in an oral stop (as in pop) and in a high-front vowel (as in poppy)? Some hypotheses were stated before data collection: for CVC words, there could be the addition of an epenthetic vowel, or possibly the palatalization, affrication or aspiration of the last consonant; for CVCV words, it was expected that learners could somehow shorten the last vowel duration, due to the fact that unstressed final vowels in BP are reduced in intensity and length and because of this are subject to vowel reduction. In order to support the hypotheses above, theories and models that could possibly explain those phonetic and phonologycal phenomena were studied and presented. First of all, the syllabic structures from English and Portuguese were contrasted to explain why Brazilian learners sometimes insert an epenthetic vowel after an oral stop in word final position to conform an unusual syllable into another that follows L1 rules. Considering that data were analyzed according to their acoustic characteristics, this research dedicated a chapter to describing how speech sound are produced based on "The Acoustic Analysis of Speech" by Kent & Read (1992). Analyzing data through the acoustic point of view was somehow new in this research field in Brazil - specifically to check whether or not Brazilian learners added an epenthetic vowel – since other studies had only counted on native speakers judgement.

One native speaker aged 16 and twelve Brazilian learners aged 13 to 17, with 250 to 300 hours of previous instruction, were recorded reading a carrier sentence containing the target words (pop/poppy, Bob/Bobby, pet/petty, dad/daddy, luck/lucky, bug/buggy). For each target word there were two distractors; the learners read three sets of sentences, organized randomly in booklets. Data were analyzed through Praat, a free software developed for acoustic analysis of speech. The results showed that, at least in an intermediate level of proficiency, students produce the epenthetic vowel

in only 7% of the CVC words. However, they do not seem to produce the final consonant of CVC targetlike: a variety of phonetic processes were found in their production. In some cases learners produced the aspirated bilabials and palatalized alveolars, but the most common phonetic characteristic was the affrication of the voiceless alveolar stop. In the CVCV words, in almost all cases learners produced shorter final vowels than the native speaker. It was not possible to confirm the tendency that Brazilian learners frequently shorten the duration of the final vowel, and that could be due to a characteristic of Portuguese language, mainly because of lack of data produced by native speakers. Processes such as aspiration, affrication and palatalization of the second consonant were also found, usually together with devoicing of the final vowel or even with its deletion.

Keywords: acoustic phonetics, vowel epenthesis, vowel reduction, duration of vowels.

The relationship between speech and singing in Bossa Nova

Ana Paula Roza – Lafape – IEL – Unicamp
anyroza@gmail.com

This honor's thesis study deals with the relationship between linguistic and musical aspects in some songs of the popular music style that came to be known as Bossa Nova (henceforth BN). Brazilian guitar player Joao Gilberto summed up the beat of the genre and developed a way of singing that was close to speech. He influenced a generation of musicians and performers (Garcia, 1999). This research investigates the relationships between BN song and speech, and music and lyrics. The parameters used were: in acoustic analysis, measures of tone, syllable and pause duration; and in categorical analysis, traditional metric units. The hypotheses are: there exist relationships between prosodic and musical structure in BN; linguistic structure influences lyrics prosody.

Three Brazilian Popular Music singers were recorded with and without musical accompaniment and while reading newspaper articles and lyrics of several songs. The reading corpus was used to normalize speech duration for each performer. Score durations were estimated from the duration measures of the performances.

We analyzed the musical scores of 'Chega de Saudade' and 'Wave' as well as the interpretations and

recitations of these songs. Music and lyrics were also investigated in scores, singing and recitation. The parameters of acoustic analysis were: tone (in Hertz); syllable duration and pause (both in milliseconds).

In pause analysis, observed/expected (O/E) ratios were applied to relate the values observed in actual singing to the ones expected from the musical scores. Thus, it was investigated the whether pauses were larger or smaller than those of the score.

The first derivative of the pitch curve was used to analyze tone. It was thus possible to investigate the behavior of tone prominences on the syllables of each sung interpretation. Syllable durations were converted to z-scores to compare the singers' production with the musical scores. That measure allows for classifying durations in long, medium or short, as compared to the musical scores.

The song texts were segmented according to traditional metrical units. Thus, the prosodic

constituents were verse and *caesura*, delimited by pauses. These procedures permit to compare scores to interpretations, as well as singing to recitation.

The rhythm (*tempo*) markings on the scores (weak or strong) were tested for association with syllable stress. This comparison paired rhythmic accompaniment information with sung syllables.

Non-parametric statistics were applied to the data. The variables were: 'tone', 'position of *caesura*', 'end of verse position', 'pause', 'tone prominence', 'duration prominence' and '*tempo*'. The comparison between interpretations of melody and rhythm of playbacks revealed that strong *tempo* tends to coincide with stressed syllables, pauses, and duration prominence. The results suggest that the typical BN beat is given by its rhythmic structure, which is implemented without distorting the lyrics.

Reference

GARCIA W. Bim Bom - A contradição sem conflitos de João Gilberto. São Paulo: Paz e Terra S.A., 1999.

Speech Timing and Tonal Alignment in Non-fluent Aphasia

Anne Zimmer-Stahl.
Anne.Zimmer-Stahl.1@city.ac.uk

The fundamental assumption underlying this study is that looking at language and speech disorders can inform our knowledge of language and speech processes in general. Explaining what can go wrong provides evidence for or against models of speech and language processing as it requires specific points of reference with respect to distinct computational components as well as the kinds of representational units present in the models. This paper looks into phenomena of speech timing in non-fluent aphasia, contrasting it with dysfluencies in speech dyspraxia and stuttering (Howell, 2007).

Several studies on intonation in non-fluent aphasia (e.g., Seddoh, 2004; Niemi, 1998) supported the “Dissociation Hypothesis”, put forward by Danly and Shapiro (1982). This hypothesis claims that intonation is dissociated from 'general speech timing', based on the finding that the production of intonation seems to be rather unaffected in non-fluent aphasia while they assume 'temporal control' to be generally disturbed.

This paper considers the Dissociation Hypothesis to be conceptually flawed and presents data that calls for a clearer distinction of the notions of fluency, prosody and articulatory implementation, their disorders, and the appropriate phonetic measurements.

It is widely recognised that, in English, tones can be aligned early or late with respect to a stressed syllable, resulting in the categorical perception of a falling or rising accent (Pierrehumbert, 1980; Ladd, 1983; Pierrehumbert & Steele, 1989). That is, pitch events are not only timed with reference to one another, but via abstract phonological representations, with reference to the segmental string. Hence, the realization of tones must be integrated into the overall timing of speech. The question arises, what exactly it is that the authors mentioned in the first paragraph meant by 'speech timing'. It appears that there is a tradition of inappropriately associating non-fluency with dysprosody and articulatory deficiencies. For example, Goodglass & Kaplan (1983) characterize speech in Broca's aphasia as dysprosodic, possibly referring to a perceptual effect that rather results from the listener's inability to perceive rhythm due to non-fluent production.

Furthermore, Gandour et al. (1984, 1989 and 1994) found a lack of temporal control in speakers with Broca's aphasia, specifically in the coordination of simultaneous articulatory gestures.

Here, it seems important to make a more consistent distinction between speech and language disorders, between agrammatism, dysprosodia and dyspraxia (disorder in phonetic implementation), and to rather allow for their co-occurrence, than to mingle their effects in a syndrome diagnosis.

Data presented in this paper shows how different measures of fluency, articulatory timing and speech rhythm (as speaking rate/rate of articulation, and voice onset timing, Laver, 1994, and the vocalic Pairwise Variability Index, Low, Grabe, & Nolan, 2000) capture disorder in different areas of language and speech impairment.

The study draws on data from two mildly non-fluent aphasic speakers, and one speaker with speech dyspraxia. Picture descriptions as well as free speech samples were analysed qualitatively for apparent production difficulties, and quantitative measures such as mentioned above were calculated. The temporal alignment of tones was elicited in two experiments that carefully controlled for segmental and prosodic context. This tonal alignment data will be discussed in relation to the other speech timing measures for each individual case. Control data was provided by seven healthy participants from the same city and age group as the aphasic speakers. The evidence from this study undermines the rationale of the Dissociation Hypothesis, as it can be shown that firstly, fluency and articulatory timing are independent dimensions of speech production, and secondly, the production of intonation involves integration at the level of articulatory planning and therefore is subject to speech timing.

References:

- Danly, M. and Shapiro, B. 1982. Speech Prosody in Broca's Aphasia, *Brain & Language* 16, 2, 171-190.

- Gandour, J. and Dardarananda, R. 1984. Prosodic Disturbance in Aphasia: Vowel Length in Thai, *Brain & Language* 23, 177-204.

- Gandour, J., Petty, S. and Dardarananda, R. 1989. Dysprosody in Broca's Aphasia: A Case Study, *Brain & Language* 37, 232-257.

- Gandour, J., Dechonkit, S., Ponglorpisit, S., and Khunadorn, F. 1994. Speech Timing at the Sentence Level in Thai after Unilateral Brain Damage, Brain & Language 46, 419-438.

- Goodglass, H. & Kaplan, E. 1983. *The Assessment of Aphasia and Related Disorders*, Lea & Febiger, Philadelphia.

- Howell, P. 2007. A model of serial order problems in fluent, stuttered and agrammatic speech, *Human Movement Science* 26, 728–741

- Ladd, D.R. 1983. Phonological Features of intonational peaks. *Language* 59, 721-759. - Laver, J.

1994. *Principles of Phonetics*. Cambridge University Press, Cambridge, UK.

- Low, E.L., Grabe, E. & Nolan, F. 2000. Quantitative Characterizations of Speech Rhythm: Syllable-Timing in Singapore English, *Language & Speech* 43, 377-401.

- Niemi, J. 1998. Modularity of Prosody: Autonomy of Phonological Quantity and Intonation in Aphasia, *Brain & Language* 61, 45-53.

- Pierrehumbert, J. 1980. *The Phonology and Phonetics of English Intonation*. PhD dissertation, MIT. Reproduced by Indiana University Linguistics Club, Bloomington.

- Pierrehumbert, J. and Steele, S. 1989. Categories of Tonal Alignment in English, *Phonetica* 46, 181-196.

- Seddoh, A.S. 2004. Prosodic Disturbance in Aphasia: Speech Timing versus Intonation Production, *Clinical Linguistics and Phonetics* 18, 1, 17-38.

Similarities and differences between sung and spoken speech.

Antônio Pessotti – Lafape – IEL – Unicamp
antoniopessotti@yahoo.com.br

Music and language have multiple interactions. They can be compared as to rhythm, intonation and phonetic-acoustic features. Speech has its own constraints, and its production is dependent on context. Singing has constraints imposed by melody and rhythm.

This doctoral research in progress attempts to describe the possible influence of musical and linguistic constraints on sung performance by means of phonetic-acoustic tools. Fifteen subjects with same voice range (soprano) were recorded. These were divided into three groups: five professional radio news announcers (non singers), five non-professional singers and five professional singers. They had to read the lyrics of Brazilian chamber song "Conselhos", by Carlos Gomes five times. All singers have sung ten times, five times with musical accompaniment and five without.

Acoustic and categorical variables were used in the analysis. The acoustic variables were: syllable and segment duration (in milliseconds, ms), pitch (in semitones, st, reference 440Hz) and F1, F2 and F3 (in Bark). The categorical variables were: stress pattern (Mattoso Câmara, 1977: 0 for posttonic, 1 for pre-tonic, 2 for non-prominent tonic, 3 for prominent tonic) and metric syllables (s - strong, w - weak). Tonal and durational prominence variables were obtained by pitch and duration first derivative curves. Tonal alignment was evaluated with *Dynamic Time Warping* (Coleman, 2005). Segmentation of recordings and data extraction were performed by semi-automatic procedures using Praat (version 5.1.23, www.praat.org) and R (version 2.9.0, www.r-project.org).

It is expected that sung text intelligibility might be related to rhythmic and melodic precision in performance. If intelligibility requires an adaptation of segmental production, it should probably be correlated to acuity. If these settings go beyond the act of singing, they could possibly revert to spoken production, resulting in a "professional accent" in speech.

In sung performance without musical accompaniment (*a capella*), descriptive statistics (mean, standard deviation, skewness, kurtosis) of durational and intonational data showed that some nonprofessional informants did not completely match melody and rhythm. Apparently, professional

singers attain the right matches easily.

Vowel spaces plotting revealed that vowel quality in professional singing remained close to that in spoken production and similar to that produced by non-singers. On the other hand, non-professional singers present several differences between spoken and sung vowels.

This may be related to articulatory degrees of freedom and indicates possible carry over from singing to speaking. The same process seems to occur in second language acquisition, in which learning is facilitated by singing practices (Slevc and Miyake, 2006; Milovanov et al, 2008). It was also observed that linguistic phrasing was influenced by musical constraints at the syllable and higher levels.

One of our questions: could a dynamic approach shed some light on these data?

References

COLEMAN, J. **Introducing speech and language processing**. Cambridge University Press, Cambridge, UK, 2005.

MATTOSO CÂMARA Jr, J. **Contribuição à Estilística Portuguesa.** Rio de Janeiro; Ao Livro Técnico; 1977.

SLEVC, L.R., MIYAKE, A. **Individual differences in second language proficiency: does musical ability matter?** Psychol. Sci. 17, 675–681; 2006.

MILOVANOV, Riia HUOTILAINEN, Minna VÄLIMÄKI, Vesa ESQUEF, Paulo A.A. TERVANIEMI, Mari **Musical aptitude and second language pronunciation skills in school-aged children: Neural and behavioral evidence**; Brain Research 1194, 81 – 89; 2008

The effect of prominence position on boundaries in Greek[1]

1 This work will be also presented at the 12th Conference of Laboratory Phonology, Albuquerque, New Mexico, 2010.

Argyro Katsika
argyro.katsika@yale.edu

This study examines how boundaries and prominence interact in temporally shaping speech production within the framework of Articulatory Phonology. Utterances undergo acoustic and articulatory durational modifications near phrase boundaries and under prominence. At positions preceding and following stronger boundaries, acoustic durations are longer (e.g. Oller 1973) and articulatory gestures are longer, larger and less overlapped (e.g. Byrd & Saltzman 1998). These boundary effects are cumulative, i.e. increasing with boundary strength (e.g. Byrd 2000; Cho & Keating 2001; Fougeron 2001). As for the prominent positions, they appear with longer acoustic durations (e.g. Turk & Sawusch 1997) and longer, faster and less overlapped gestures (e.g. Cho 2006). The spatiotemporal boundary effects on articulation have been modeled by a model integrated within Articulatory Phonology, called π-gesture (Byrd & Saltzman 2003). According to this model, boundaries are prosodic gestures that extend over temporal intervals, during which they slow down the constriction gestures that are co-active with them.

A theoretically and empirically open question is the scope of effect of these boundaries, specifically how it relates to prominence. While articulatory (e.g. Byrd, Krivokapić & Lee 2006, Krivokapić 2007) and acoustic (e.g. Berkovits 1993a, b, 1994; Shattuck-Hufnagel & Turk 1998) studies have shown that boundary effects might extend away from the boundary, the scope of the speech interval affected by the boundary and its relation to the locus of prominence are unclear. To date, the only two studies examining this question show different effects. An acoustic study by Turk & Shattuck-Hufnagel (2007) has found that the boundary effect scopes over material local to the boundary and the prominent syllable, making the domain of the effect discontinuous when the prominent syllable is away from the boundary. However, Byrd and Riggs (2008) observe in an articulatory study that boundaries affect a continuous domain local to the boundary (as predicted by the π -gesture model), which, for some speakers, extends to the nearby prominent syllable. A further open question is whether the results obtained from these two studies are specific to English and whether different languages might show a different pattern of prosodic behavior.

The current study examines the temporal scope of boundaries of different strength as a function of prominent position both pre- and post-boundary in Greek. Two acoustic experiments

were conducted manipulating boundary type (word, minor phrase and major phrase) and the distance of the prominent syllable from the boundary (boundary-adjacent, 1 or 2 syllables away from the boundary). The two experiments manipulated the two factors pre- and post-boundary respectively, yielding 18 test sentences. To control for segmental effects, the different prominence positions were acquired by the means of the neologisms *MAmama*, *maMAma* and *mamaMA* (capital letters indicate stress). Data from 12 repetitions of each of the 18 sentences produced by 5 native Greek speakers have been collected. The effects of boundary and prominence, and their interaction, on the acoustic duration of the boundary pauses and the test word's segments have been statistically evaluated.

The results indicate speaker-specificity in determining the scope of the boundary effect: two speakers had continuous scope, two speakers had discontinuous scope and one speaker had continuous scope pre-boundary and discontinuous post-boundary. All speakers showed an effect on the boundary-adjacent material. The speakers with discontinuous domain also showed an effect on either the prominent syllable or (when post-boundary) the other edge of the word. Segmental modifications were equal or larger in minor phrases than in major phrases, while the pauses of the former, when present, were much shorter than the pauses of the latter.

These results are discussed within the Articulatory Phonology framework. It is suggested that minimally boundaries affect immediately adjacent material, with more distant effects being speaker-specific. The effect is cumulative in terms of pausal durations, but not in terms of segmental durations. The π-gesture model captures the general pattern of speakers' productions in terms of the use of pauses and segment durations. We discuss the speaker-specific patterns of π-gesture coordination, and how the relationship between acoustics and articulation plays out in our results.

References

Berkovits, R. (1993a). Progressive utterance-final lengthening in syllables with final fricatives. Language and Speech 36 (1), 89–98.

Berkovits, R. (1993b). Utterance-final lengthening and the duration of final-stop closures. Journal of Phonetics 21, 479–489.

Berkovits, R. (1994). Durational effects in final lengthening, gapping, and contrastive stress. Language and Speech 37 (3), 237–250.

Byrd, D. (2000). Articulatory vowel lengthening and coordination at phrasal junctures. Phonetica 57, 3-16.

Byrd, D. & Riggs, D. (2008). Locality interactions with prominence in determining the scope of

phrasal lengthening. Journal of the International Phonetic Association 38 (2), 187- 202.

Byrd, D. & E. Saltzman (1998). Intragestural dynamics of multiple phrasal boundaries. *Journal of Phonetics*, 26: 173-199.

Byrd, D. & Saltzman, E. (2003). The elastic phrase: Modeling the dynamics of boundaryadjacent lengthening. Journal of Phonetics 31, 149-180.

Byrd, D., Krivokapić, J. & Lee, S. (2006) How far, how long: On the temporal scope of prosodic boundary effects. Journal of Acoustical Society of America 120 (3), 1589- 1599.

Cho, T. (2006). Manifestation of prosodic structure in articulatory variation: Evidence from lip kinematics in English. In: Laboratory Phonology 8: Varieties of Phonological Competence. Edited by L. Goldstein. Berlin/New York: Mouton De Gruyter, pp. 519- 548.

Cho, T. & Keating, P. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. Journal of Phonetics 29, 155-190.

Fougeron, C. (2001). Articulatory properties of initial segments in several prosodic constituents in French. Journal of Phonetics, 29, 109-135.

Krivokapić, J. (2007). The planning, production, and perception of prosodic structure. Unpublished Ph.D. thesis, University of Southern California.

Oller, D. K. (1973). The effect of position in utterance on speech segment duration in English, Journal of the Acoustical Society of America 54 (5), 1235-1247.

Shattuck-Hufnagel, S. & Turk, A. (1998). The domain of final-lengthening in English. Journal of the Acoustical Society of America 102 (2), 1235–1236.

Turk, A. & Sawusch, J. R. (1997). The domain of accentual lengthening in American English. Journal of Phonetics 25, 25-41.

Turk, A. & Shattuck-Hufnagel, S. (2007). Multiple targts of phrase-final lengthening in American English words. Journal of Phonetics 35 (4), 445–472.

How /b, d, g/ differ from /p, t, k/ in Spanish: A dynamic account

Benjamin Parrell
parrell@usc.edu

The traditional view of /b d g/ in Spanish is that these segments are underlying stops, contrasting in voicing with /p t k/, that undergo a process of spirantization except phrase initially and after a homorganic nasal or lateral (e.g. Hualde, 1988; Mascaró, 1984).

However, a large number of authors argue that these segments are underlying *spirants*, contrasting with /p t k/ in constriction degree, that undergo initial and post-nasal strengthening (e.g. Baković, 1995; Lavoie, 2001). Both of these theories rely on categorical productions of fully occluded versus fully spirantized stops. A large body of evidence shows that this is not the case in real speech: voiced stops in all contexts span a large area between these two extremes (Carrasco & Hualde, 2009; Cole, Hualde & Iskarous, 1999; Eddington, 2009; Ortega-Llebaria, 2004). Additionally, voiceless stops can be realized both as fully voiced and spirantized (Lewis, 2001; Machuca, 1997). How, then, are /b d g/ and /p t k/ different? Some authors have proposed that duration (Hualde, 2005) or the feature [±tense] (Martínez Celdrán, 2008 and previous) distinguish /b d g/ from /p t k/.

This paper argues for a unified analysis of Spanish stop spirantization and contrast based in Articulatory Phonology (Browman & Goldstein, 1992). The shorter duration of voiced stops (e.g. Lavoie, 2001) may lead to their less constricted productions (spirantization) due to increased articulatory undershoot. It is that duration difference, in addition to voicing, that distinguishes /p t k/ from /b d g/. Increased duration phrase initially and due to nasal + stop sequences lead to productions of full stops in those positions without overt allophonic control.

Methods: To test this hypothesis, a pilot study was conducted using 3d electromagnetic articulometry (EMA) with one native speaker of northern peninsular Spanish. The subject produced words in carrier phrases with /p b t d/ in /a(#)Ca/ context phrase initially, word initially, and word medially. The duration of the gesture from onset to constriction release for all tokens was measured, along with the constriction degree (CD, distance between sensors on upper and lower lips) for labial stops (positioning of the tongue tip sensor, ~ 8 mm behind the apex, resulted in inaccurate constriction measurements for coronals).

Result: For total duration, there is a main effect of prosodic boundary ($p < 0.0001$) and a

near-significant effect of voicing ($p = 0.08$). T-tests reveal significant differences both wordinitially and medially ($p < 0.0001$, 0.03). For CD, there were main effects of both prosodic boundary ($p = 0.0002$) and voicing ($p < 0.0001$). Post hoc test reveals a significant difference between phrase medial /b/, and phrase initial /b/ and /p/ in all prosodic positions ($p < 0.0001$).

Linear regressions (Figure 1) show a significant effect of duration on CD for /b/ ($p < 0.005$), though not for /p/. As can be seen from Figure 1, however, /b/ and /p/ with the same duration differ in CD, against our initial hypothesis. An ANOVA test reveals no difference for gestural stiffness, ruling out that parameter as the cause of this difference. However, it is known that stops have a CD target beyond the point of articulator contact (Löfqvist, 2005).

Though /b/ and /p/ do differ in CD, it is proposed that /b d g/ have a *less negative* target than /p t k/ but still one that results in full occlusion. An articulatory modeling study conducted using TaDA (Saltzman, Nam, Krivokapić & Goldstein, 2008), confirms that a CD for /p t k/ of -2 mm and for /b d g/ of -0.5 mm gives the correct articulatory and acoustic output both phrase-medially and initially (duration = 80 ms, 200ms respectively) while a CD for /b d g/ of 0 mm results in incomplete closure and audible frication even at long durations (Figure 2).

In summary, these data indicate that a) voiced stops in Spanish are, in fact, stops and not spirants and b) the distinction between voiced and voiceless stops must be at least two-way, with differences in duration and constriction degree in addition to a possible voicing distinction.

Figure 1

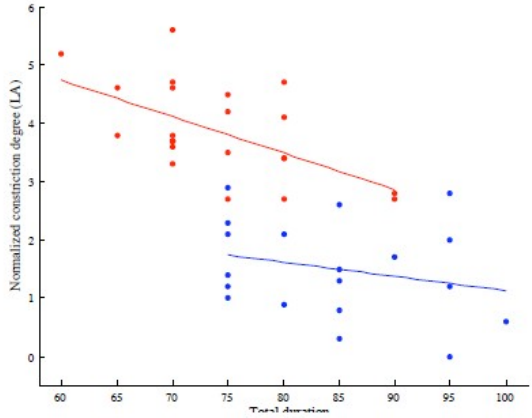
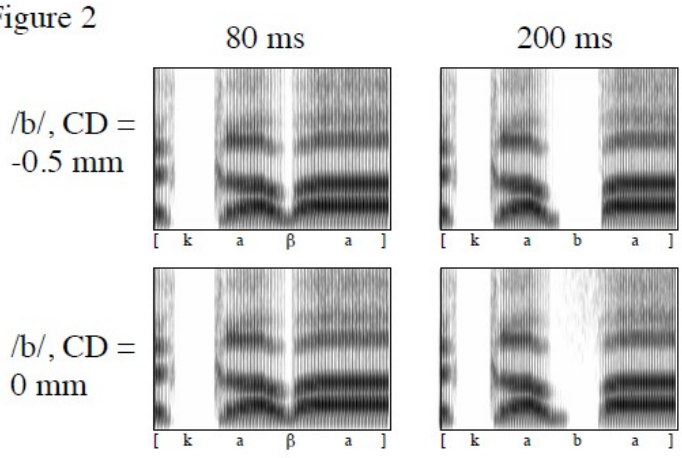


Figure 2



References:

Baković, E. (1995). Strong onsets and Spanish fortition. *MIT Working Papers in Linguistics,Proceedings of SCIL 6*, 21-39.

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, *49*(3-4), 155-80.

Carrasco, & Hualde, J. I. (2009). *Spanish voiced allophony reconsidered*. Paper presented at the Phonetics and Phonology in Iberia Conference, Las Palmas, Spain.

Cole, J., Hualde, J., & Iskarous, K. (1999). Effects of prosodic and segmental context on /g/-lenition in Spanish. In O. Fujimura, B. D. Joseph, & B. Palek (Eds.), *Proceedings of the fourth international linguistics and phonetics conference.*

Eddington, D. (2009). *What are the contextual phonetic variants of /β, d, γ/ in colloquial Spanish?* Unpublished manuscript, Brigham Young University.

Hualde, J. I. (1988). *A lexical phonology of Basque.* Ph.D. Dissertation, USC.

Hualde, J. I. (2005). *The sounds of Spanish.* Cambridge: Cambridge University Press.

Lavoie, L. M. (2001). *Consonant strength: Phonological patterns and phonetic manifestations.* New York: Garland.

Lewis, A. M. (2001). *Weakening of intervocalic /p, t, k/ in two Spanish dialects : Toward the quantification of lenition processes.* Ph.D. Dissertation, UI-UC.

Löfqvist, A. (2005). Lip kinematics in long and short stop and fricative consonants. *Journal of the Acoustical Society of America*, *117*(2), 858-878.

Machuca, M. J. (1997). *Las obstruyentes no continuas del español: Relación entre las categorías fonéticas y fonológicas en habla espontánea.* Ph.D. Dissertation, Universitat Autònoma de Barcelona.

Martínez Celdrán, E. (2008). Some chimeras of traditional Spanish phonetics. In L. Colantoni, & J. Steele (Eds.), *Selected proceedings of the 3rd conference on laboratory approaches to Spanish phonology.* (pp. 32-46). Somerville, MA: Cascadilla.

Mascaró, J. (1984). Continuant spreading in Basque, Catalan and Spanish. In M. Aronoff, & R. Oehrle (Eds.), *Language sound structure.* (287-98). Cambridge, MA: MIT Press.

Ortega-Llebaria, M. (2004). Interplay between phonetic and inventory constraints in the degree of spirantization of voiced stops: Comparing intervocalic /b/ and intervocalic /g/ in Spanish and English. In T. L. Face (Ed.), *Laboratory approach to Spanish Phonoloogy.* (237-54). Berlin: Mouton de Gruyter.

Saltzman, E., Nam, H., Krivokapić, J., & Goldstein, L. (2008). A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. *Proceedings of the Speech Prosidy 2008 Conference*, 175-84.

ADoReVA and ADoTeVA, Two PRAAT Plugins for the Automatic Detection of Register and Tempo Variations

Celine de Looze.
celine.delooze@gmail.com

As pointed out many years ago by Bolinger [4], a major drawback of most scalar systems for representing intonation patterns is the difficulty in separating global pitch changes (determined by variations in register key and span) from local pitch characteristics (determined by changes in the phonological representation of intonation). How can we distinguish for example a high fall in a narrow register from a low fall in a wide register? To answer this argument, changes in the f0 domain are accounted for by admitting only two level tones as assumed in AM theory ([17, 18, 19]) or more, as with INTSINT ([14]). However, while these models appear adequate for the analysis of short read sentences (as often employed in laboratory speech), the fact that they implicitly assume that a speaker's key and span remain unchanged makes their use fragile for the analysis of spontaneous speech, where variations in register may convey, among other things, information about the speaker's identity or about the discourse structure as well as the hierarchical dimension and relational organisation of the discourse. The same is true for the temporal organisation of speech.

Models have to cope with the overlap between, on the one hand, short-term features such as segmental duration and longer-term ones such as tempo variations which convey extralinguistic as well as linguistic functions and which therefore make the analysis and modelling of the temporal organisation of speech difficult. How can we distinguish for example a short sound in a slow tempo from a long sound in a fast tempo? These difficulties show the importance of understanding the temporal span of longer-term prosodic variations when describing short-term ones especially when one is interested in the study of spontaneous speech and its functions.

The difficulty in defining the temporal span of register and tempo variations comes from the fact that these variations operate over many different domains. We consequently propose two clustering algorithms, ADoReVA and ADoTeVA, which automatically detect variations in register and tempo respectively. First of all, the algorithms calculate the difference in register and tempo between two consecutive units. They then generate a binary tree structure in the form of a layered icicle diagram which provides the graphical representation of register (Figure 1) and tempo

variations. This representation defines the hierarchical structure and relational organisation of discourse units as reflected by changes in register and tempo.

Groups of units are therefore distinguished and the distance between the leaf nodes provides a measurement of boundary strength between them. The greater the distance, the stronger the boundary between two groups, while a short distance suggests that two consecutive units belong to the same group of units.

We have shown for example, using the Aix-Marsec corpus [1], (a corpus of spontaneous British English with a total of 5h30 of speech) that the units at the boundary of two major intonation units were marked by a larger difference in register and articulation rate than the units at the boundary of two minor intonation units or within minor intonation units. These results suggest that one should take into account the hierarchical and embedded-domains structure of register and tempo variations when studying these long-term phenomena as well as when studying shorter term prosodic features.

In addition, we applied register variations such as obtained with the algorithm ADoReVA to the MOMEL-INTSINT algorithm ([14]) – which allows the automatic coding of intonation patterns – and showed that integrating these variations significantly improves the automatic coding of intonation patterns.

Finally, an automatic detection of register and tempo variations leads to a better understanding of the functions they convey, such as the way they indicate the speaker's identity or the intentional structure of the discourse. In our study, we have shown, using four corpora (the CID [3], the PFC [5] for the study of French and the AIX-MARSEC and the PAC corpora [5], for the study of English – a total of 7h30 speech), that (1) French and English were not distinguished by differences in register and tempo despite their different prosodic structure, that (2) men and women were differentiated by register only but not by tempo and that (3) different styles of speech were indicated by different tempi only (articulation rate and pause duration). We also showed a strong correlation between discourse structure and register and tempo variations, where topic changes are indicated with higher and expanded register and with longer pauses and variations in articulation rate.

It is consequently important to take into account variations in register and tempo, especially when studying spontaneous speech. The difficulty in defining their temporal span is partly resolved by the use of the two algorithms we presented, although further study is needed to better understand the functions these variations convey. We also suggest that these variations could be used with existing tools for the automatic coding of intonation patterns or the temporal organisation of speech in order to better describe short-term prosodic variations.

Figures

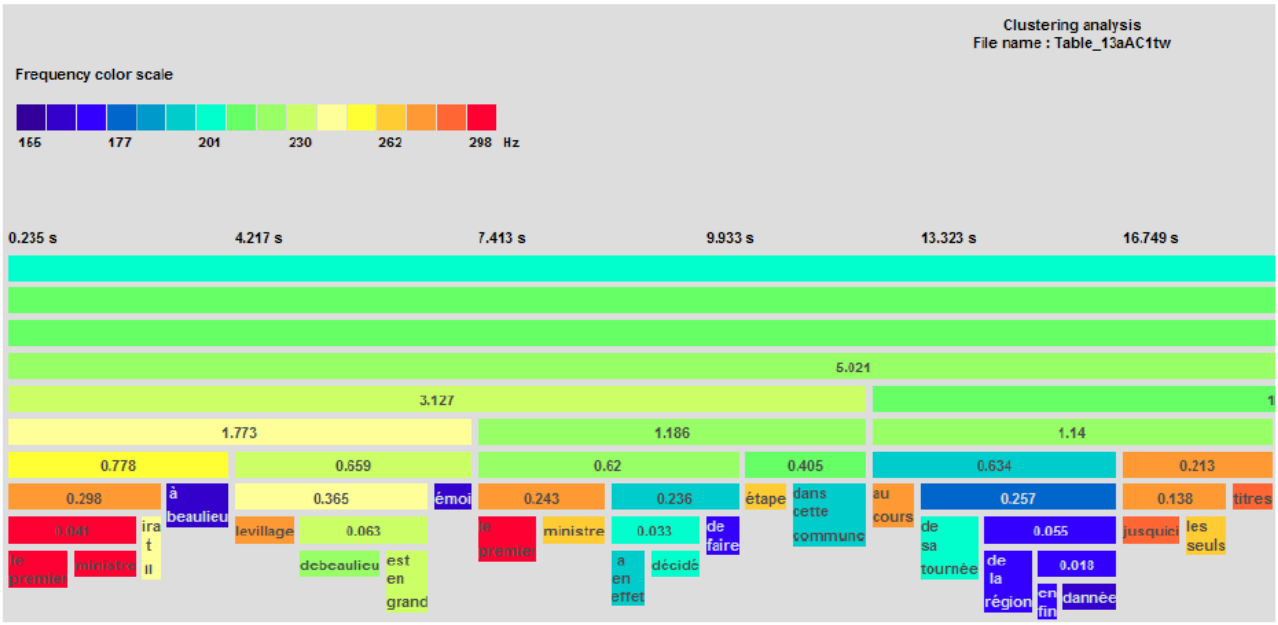


Figure 1. Extract of a layered icicle diagram representation as obtained with the algorithm ADoReVA. Units are grouped together according to their register level and span. The representation shows that the unit "le premier" and "ministre" belong to the same group (according to the register level of each unit); on the contrary, we clearly distinguish the group of units "le premier ministre ira-til à Beaulieu" from the group of units "le village de Beaulieu est en grand émoi". In fact, the distance between the leaf nodes "à Beaulieu" and "le village" indicates a strong frontier (corresponding to a break in the tree structure). The colour scale indicates the register level for each unit. The warmer the colour, the higher the key.

References

[1] Auran, C., Bouzon, C. & Hirst, D.J. (2004) The Aix-MARSEC Project: An Evolutive Database of Spoken British English, in *Speech Prosody* 2004, ISCA.

[3] Bertrand, R., Blache, P., Espesser, R., Ferré, G., Meunier, C., Priego-Valverde, B. and Rauzy, S. (2007) Le CID-Corpus of Interactional Data-: protocoles, conventions, annotations in *Travaux Interdisciplinaires du Laboratoire Parole et Langage d'Aix* en Provence, 25, 25- 55.

[4] Bolinger, D. (1951) Intonation: Levels vs. Configurations, in *Word,* 7, 199-210.

[5] Delais-Roussarie, E. & Durand, J. (2003) *Corpus et variation en phonologie du français: méthodes et analyses*, Presses Universitaires du Mirail.

[14] Hirst, D.J. (2007) A Praat Plugin for MOMEL and INTSINT with improved algorithms for modelling and coding of intonation*,* in *Proc. Int. Conf. Phonetic Sci. XVI,* Saarbrucken.

[17] Ladd, DR. (1996) *Intonational Phonology,* Cambridge University Press, Cambridge.

[18] Pierrehumbert, J. (1980) *The Phonology and Phonetics of English Intonation*, PhD Dissertation, Cambridge, Mass., MIT.

[19] Silverman, K. et al (1992) ToBI: A Standard for Labeling English Prosody in *Second International Conference on Spoken Language processing.* ISCA.

The effects of Lombard speech on vowel formant measurements

Christin Kirchhuebel
ckirchhuebel@gmail.com

Introduction

It is well known that people tend to increase their vocal effort when speaking on the phone or against loud background noise which has come to be known as the Lombard reflex. Knowing the auditory and acoustic effects of 'Lombard speech' i.e. speech produced under noise, is of important practical relevance to forensic phonetics as well as speech science in general.

Forensic speaker comparison cases involve comparing a questioned sample and a known sample of speech in order to assess the likelihood that both have been produced by the same speaker. In the majority of cases the disputed recording has been transmitted over the telephone (Byrne & Foulkes 2004:83 quoting Hirson French & Howard 1995:230) and consequently it can be assumed, that a great number of speaker comparison cases involve comparing modal speech with Lombard speech. On the other hand, scientist working on automatic speech recognition may also benefit from research on the acoustic effects of Lombard speech. In order for the systems to perform well they need to be robust to intraspeaker variability arising from internal factors such as speaker state but also external factors such as background noise.

Background literature

There are a range of studies looking at the effects of Lombard speech on acoustic parameters such as fundamental frequency, vowel duration, intensity and to a lesser extent vowel formants. (Harris 1963, Summers et al. 1988) Based on the above studies, Junqua (1996) summarises the acoustic characteristics associated with the Lombard reflex in the following way:

- increase in f0
- increase in energy from low frequency bands to middle or high bands
- increase in level
- increase in vowel duration
- spectral tilting
- shift in formant centre frequencies for F1(mainly) and F2 (Junqua 1996:16)

These studies have also shown that 'the Lombard reflex is highly variable from speaker to

speaker' (Junqua 1996:16). More recently, researchers have become interested in this speaker dependency and started investigation into the inter- and intra- speaker variability between modal and Lombard speech with regard to f0 (Jessen & Köster & Gfroerer 2005, Beckford Wassink et al. 2006), hesitation markers (Tschäpe et al. 2005), segment duration and F1/F2 frequencies (Beckford Wassink et al. 2006).

Aim of this study

This present study intended to analyse the acoustic effects of Lombard speech on the vowel formants F1, F2 and F3. In particular, attention was given to inter-speaker variability and attempts have been made to make generalisations from the observed patterns.

Methodology

Ten male native German speakers were selected from the 'Pool 2010' corpus which was recorded at the Bundeskriminalamt (BKA), Germany. Spontaneous speech produced in a neutral setting and Lombard setting, where 80dB of noise was played through headphones, was analysed. Measurements of F1, F2 and F3 were collected from 10 vowel categories for every speaker in both conditions. The Speech Analysis Software 'Praat' was used to analyse and log the data.

Findings

The results agree with previous findings in that F1 is consistently higher in the Lombard condition. The effect on F2 is very variable and complex. F3 was less affected than F1 and F2, but changes were present, especially for speakers with low F3s in modal speech. Differences could be observed among vowel categories. Inter-speaker variability was found to be large with respect to the size of increase in F1 and the direction and size of change in F2. The findings are discussed in light of the articulatory changes that have been associated with Lombard speech and the implications for forensic speaker comparison and automatic speech recognition are spelled out.

keywords: forensic speaker comparison, Lombard speech, vowel formants, inter-speaker variability, automatic speech recognition

References

1. Beckford Wassink, A. & Wright, R.A. & Franklin A. D. (2006). Intraspeaker variability in vowel production: An investigation of motherese, hyperspeech, and Lombard speech in Jamaican speakers. *Journal of Phonetics*: 35 (2007) 363–379

2. Byrne, C. & Foulkes, P. (2004). The mobile phone effect on vowel formants. *The International Journal of Speech, language and the Law.* 11: 83-102.

3. Harris, C. M. (1963). Pitch and formant shifts accompanying changes in speech power level.

Journal of the Acoustical Society of America.35: 1876.

4. Jessen, M. & Köster, O. & Gfroerer, S. (2005). Influence of vocal effort on average and variability of fundamental frequency. *The International Journal of Speech, language and the Law* 12(2): 174- 213

5. Junqua, J. (1996). The influence of acoustics on speech production: A noise-induced stress phenomenon known as the Lombard reflex. *Speech Communication* 20: 13-22

6. Künzel, H.J. (2001). Beware of the 'telephone effect': the influence of telephone transmission on the measurement of formant frequencies. Forensic Linguistics: *The international Journal of Speech, Language and the Law.* 8: 80-99.

7. Summers, W. Van, Pisoni, D. B., Bernacki, R. H., Pedlow, R. I. and Stokes, M. A. (1988). Effects of noise on speech production: Acoustic and perceptual analyses. *Journal of the Acoustical Society of America*. 84(3): 917.28.

8. Tschäpe, N. & Trouvain, J. & Bauer, D. & Jessen, M. (2005). A study of the idiosyncratic patterns of filled pauses in German. Paper presented at the IAFPA conference, Marrakesh. *The international Journal of Speech, Language and the Law.* 12(2): 288

Aspiration in Scottish Gaelic stop consonants

Claire Nance – University of Glasgow
c.nance.1@research.gla.ac.uk

Introduction

Scottish Gaelic has two series of stops, orthographically <p t c> and <b d g>. These are realised in word initial position as voiceless aspirated /p^h t^h k^h/ and voiceless unaspirated /p t k/ respectively. In word medial and final position the <p t c> series are voiceless preaspirated /^{h}p ^{h}t ^{h}k/ (Ladefoged et al. 1998). Stops can also be either velarised or palatalised, but this study investigates the velarised set only.

Here preaspiration is defined, as Gordeeva and Scobbie (2010), as the presence of glottal frication. This therefore includes breathy voicing in the vowel and voiceless frication of the type described in classical preaspiration studies such as Thráinsson (1978). Phonological literature on preaspiration such as Thráinsson (1978) has considered preaspiration within a segmental or autosegmental framework. This does not adequately account for the Gaelic data, and as suggested by Casserly (2010), a gesture-based account focussing on the timing of different components to preaspiration may be more appropriate. Phonetic descriptions support this view such as Laver (1994:356) who describes preaspiration as a 'coordinatory relationship between a vowel and a following voiceless segment', and Hansson (2003:69):

'The implementation of preaspiration unavoidably involves articulatory overlap with the preceding vowel… The glottal opening gesture associated with the stop is initiated 'prematurely', and part of the preceding vowel is thus 'taken over' by preaspiration.'

Previous phonetic studies of preaspirated consonants such as Ní Chasaide (1985), Jones and Llamas (2003), Helgason (2002) have only measured durations without providing a measure of the noisiness of preaspiration. Studies have disagreed on what exactly to measure, and have categorically included preaspiration as a part of either the preceding vowel or the following consonant. This does not seem logical in the light of Laver and Hansson's comments.

Research questions

This study examines the Gaelic stop series from three different perspectives: [1] Are the contrasts between <p t c> and <b d g> maintained as previously reported i.e. by the presence or

absence or aspiration in word initial position, and by the presence/absence of preaspiration in word medial and final position? [2] Is this system changing as the language undergoes obsolescence (Dorian 1981)? [3] What framework can adequately account for the preaspirated stops?

Methods

Word list data from six native speakers of Lewis (Outer Hebrides) Gaelic were recorded in a noise-attenuated sound studio. Three were aged 40-55, and three aged 20-24, five female and one male. The inclusion of an age-graded sample allowed the investigation of a possible apparent time change. Target words were couched in a carrier phrase and grouped semantically. Each word was elicited three times. The resulting data contain examples of word initial, medial and final stops at all (velarised) places of articulation.

Durational measures (total vowel, modal voicing, breathy voicing, total preaspiration, voiceless frication, stop closure, VOT) were taken off the waveform in Praat. This comprehensive approach to durations across the syllables allows for a broader understanding of the stop in context and how preaspiration overlaps with all other components. In order to measure the noisiness of preaspiration the data were submitted to adapted Zero Crossing Rate (Gordeeva and Scobbie 2010). This is a periodicity-independent measure allowing a quantification of noise in both the breathy voiced sections and the voiceless sections of preaspiration. For comparison, adapted Zero Crossing Rate was also applied to word initial stops to examine voiceless aspiration and any following breathy voiced transitions into the vowel. The results were analysed using non-parametric statistical tests.

Results

Results from the durational measures indicate that as previously reported the contrast between the two series of stops in Gaelic is maintained through presence/absence of aspiration in word initial position, and presence/absence of preaspiration in word medial and final position.

However the system appears to be undergoing apparent time shift of the gradient phonetic type described by Babel (2009:23). Preaspiration in younger speakers is significantly shorter and contains less voiceless frication than that of older speakers. Preliminary results from the adapted Zero Crossing Rate measures suggest that this is an extremely useful method for quantifying noise in the signal relevant to both VOT measures and preaspirated sections.

Discussion

Unlike previous studies, this study provides a more complete analysis of aspiration and

preaspiration as features of the syllable, rather than assigning them as components of either the vowel or the consonant. This study therefore sides with Casserly (2010) in investigating the value of a gesture-based explanation of how to account for preaspiration.

References

Babel, M., 2009. The phonetic and phonological effects of obsolescence in Northern Paiute. In J. Stanford and D. Preston, eds., *Variation in indigenous minority languages.* Amsterdam: John Benjamins.

Casserly, E., 2010. Spread glottis in Faroese: Realization, neutralization and representation. Oral paper at the Old World Conference in Phonology, 7. Nice, France.

Dorian, N., 1981. *Language death: The life cycle of a Scottish Gaelic dialect.* Philadelphia: University of Pennsylvania Press.

Gordeeva, O., and Scobbie, J., forthcoming 2010. Preaspiration as a correlate of word-final voice in Scottish English fricatives.

Hansson, G., 2003. Laryngeal licensing and laryngeal neutralization in Faroese and Icelandic. *Nordic Journal of Linguistics,* 26:1, pp. 45–79.

Helgason, P., 2002. *Preaspiration in the Nordic languages: synchronic and diachronic perspectives*. PhD thesis, Stockholm University.

Jones, M., and Llamas, C., 2003. Fricated pre-aspirated /t/ in Middlesborough English: an acoustic study. *Proceedings of the XVth International Congress of the Phonetic Sciences.*

Ladefoged, P., and Ladefoged, J., 1998. Phonetic structures of Scottish Gaelic. *Journal of the International Phonetics Association* 28, pp. 1-41.

Laver, J., 1994. *Principles of phonetics.* Cambridge: Cambridge University Press.

Ní Chasaide, A., 1985. *Preaspiration in phonological stop contrasts: an instrumental phonetic study.* PhD thesis, University of Bangor.

Thráinsson, H., 1978. On the phonology of Icelandic pre-aspiration. *Nordic Journal of Linguistics,* 1, pp. 3-54.

An Introductory Analysis of the Roles of Duration, Intensity, and F0 in Explicit and Implicit Performatives in Brazilian Portuguese

Clerton Luiz Felix Barboza.
clertonluiz@gmail.com

Discussions about the construction of the Speech Acts Theory, as proposed by Austin (1962;1990), begin by emphasizing the difficulties in searching the distinctions among questions, orders, and statements impose to philosophers and grammarians. Such difficulties occur mainly owing to the "little and incipient grammatical criteria available" (id. Ibid.,p. 22) for the search of such a classification.

Having as starting point the search for possible grammatical distinctions between performative and constative verbs, Austin chooses to focus on the question of performative utterances, sentences which, by definition, neither describe nor declare what is being practiced. Such sentences, when produced, *do* something. These are utterances in which *by* or *while* we say something, we are actually *doing* something.

While developing the fundamental precepts of the Speech Acts Theory, Austin concludes that the performative/constative verb dichotomy is largely useless, once such a distinction does not hold against the argument in favor of "more general families of related and partially overlaid speech acts" (id. Ibid., p. 122). However, one distinction still holds at the level of performative utterances: the difference between explicit and implicit performatives. Such a distinction survives the transition from the early discussion about the differences between performatives and constatives in the Speech Acts Theory because of the existence of certain verbs (promise, proclaim, condemn, etc.) which possess a huge *illocutionary force,* helping to make explicit the speech act which is being realized at the moment.

Such an illocutionary force is understood as a speech act that realizes or tends to realize the nominated action, being promptly associated with performatives. Besides the *illocutionary* acts, there are also the *perlocutionary*, which would not be directly associated with the utterance, but to the obtained effect of what is said, and the *locutionary* ones, the acts of saying something, producing words possessing sense and reference. As concerns this last act, Austin (1962; 1990) suggests its study to the grammarians and phoneticians. Such *Speech Acts* constitute Austin's core

theory.

The fundamental question of the present research involves the idea that studies in phonetics and phonology should be used in the search of a better understanding of the production of the locutionary act of explicit and implicit performatives as proposed by Austin (1962; 1990). Such a belief is justified mainly due to fact the author (remember, a philosopher!) does not focus on questions of phonological analysis of his performative utterances. He limits himself to mention briefly the importance of certain linguistic resources, i.e., connectives, modals, adverbs, para-linguistic and circumstantial elements, as well as suprasegmental aspects such as tone of voice, cadence and emphasis for the realization of explicit performatives. A study focused on such characteristics of the utterance might help understanding the distinction between the production of explicit and implicit performatives in Brazilian Portuguese.

The main objective of this introductory research is, therefore, to describe and compare the characteristics of duration, fundamental frequency (F0) and intensity of explicit and implicit performative utterances. We aim at determining, in an interface study involving pragmatic and phonetic/phonological aspects, which of these performatives is normally associated with a higher level of prominence in Brazilian Portuguese.

Having in mind the aforementioned objective, we present as hypothesis the idea that explicit performatives are produced, owing to their characteristic of making clear their illocutionary force, with a higher degree of prominence in Brazilian Portuguese. Such hypothesis implies directly that explicit performatives will present significant higher values of duration, intensity and fundamental frequency when compared to implicit ones.

As regards our methodology, data collection procedures involved the recording of five performatives, all used in both explicit and implicit forms. We analyzed the performatives *condenar* (condemn), *proclamar* (proclaim), *prometer* (promiss), *agradecer* (thank), and *afirmar* (affirm). Explicit performatives were produced in first person sentences, having penultimate-syllable stress. Implicit performatives in third person sentences, having lastsyllable stress. F0 and intensity measures were carried out at the accented vowels. Duration was collected taking in consideration the whole word.

Informers were selected among university students. Ten were chosen, ranging from 19 to 32 years. All of them were from cities nearby Mossoró-Rio Grande do Norte-Brazil. Having in mind the number of informants, we had a total of 300 analyzed verbs, two groups of 150 explicit and implicit performatives. With each verb contributing with a value of duration, intensity and F0, the total number of variables was as high as 900.

Hardware used for data collection was an M-Audio external sound card model MobilePre

USB and a Shure headset microphone WH20XLR. Software used involved Audacity version 1.3.5, for sound editing, and Praat version 5.1.19, for acoustical analyses. Statistical analyses involving the variables were performed with SPSS version 16.

A superficial analysis of the data points out that explicit performatives were mostly realized with higher duration, F0 and intensity values than implicit ones. In three moments only did this fact not occur: duration of the verbs *prometer, agradecer* and *afirmar.* Besides, observed data was rather homogeneous once closely-related values of the mean, median, and small standard deviation, were the norm.

Results of 15 paired-samples t-tests comparing duration, F0 and intensity values of the aforementioned verbs in both explicit and implicit roles indicated 9 significant differences. As regards the verb *condenar,* the explicit performative was realized with higher variable values than its implicit counterpart. On the other hand, the explicit forms of the verbs *proclamar, prometer* and *afirmar* were significantly higher than their implicit form in F0 and intensity variables only, as no difference was found in the duration variable. Finally, the explicit and implicit forms of the performative verb *agradecer* showed no significant difference in any of the variables analyzed.

Having these results in mind, we noticed 80% of the intensity and F0 measures were significantly different in our statistical tests. Such values indicated acoustic parameters were used in order to give the accented syllable of explicit performatives peculiar characteristics in Brazilian Portuguese. Total duration of explicit and implicit performatives resulted, however, in significant differences in 20% of the data only, indicating this acoustic parameter was not highly productive to distinguish the different forms of performative verbs.

Our hypothesis was, thus, *partially confirmed.* The absence of significant differences in most duration comparisons between explicit and implicit performatives, as well as the fact no difference was found in the realization of the verb *agradecer*, was the key for this conclusion. However, the intensity and F0 tendencies found here indicate research on performatives might greatly benefit from a methodology based on acoustic studies.

Word-final devoicing in Russian and German: Electroglottographical Analysis

Darja Appelganz
darjaappelganz@googlemail.com

The phonetic and phonological status of final devoicing of obstruents has been for long time a controversial topic. The traditional line of research considers that final devoicing produces absolute neutralization of the feature [voice] in word- and/or phrase-final position in several languages e.g. Russian, German, Dutch and Polish, among others. Consider the examples below from Russian (Halle 1971, Jassem & Richter 1985, Kenstowicz and Kissenberth 1979).

/trut/ -> [trut] 'to rub, 3Pl.'

/trud/ -> [trut] 'labour, work'

However, alternative studies have shown that neutralization is incomplete so that there are subphonemic differences between voiced and voiceless consonants (Charles-Luce 1985, Charles-Luce & Dinnsen 1983, Dmitrieva & Jongmann 2005,Kanibolotskaia 2008, Port & O'Dell 1985, Warner, Jongmann, Sereno & Kepms, 2001).

While the previous research on this topic has been based on acoustic analysis (Charles-Luce 1985, Charles-Luce & Dinnsen 1983, Dmitrieva & Jongmann 2005,Kanibolotskaia 2008, Port & O'Dell 1985, Warner, Jongmann, Sereno & Kepms, 2001), there are few instances of studies addressing the issue using more direct methodologies to assess the activity of the vocal folds (Smith, 1997). In my B.A. Thesis I present study which is based on an electroglottographic (EGG) and acoustic analysis of the final devoicing in Russian and German. The combination of these approaches will yield a more comprehensive account of the phenomenon under investigation.

The research question guiding the investigation was whether the voiceless/voiced distinction in final position was or was not neutralized in Russian and German, and whether these languages behave in the same way. Furthermore, because previous reports suggest that neutralization depends on the level of phrasal position, i.e. medial vs. final, I designed a paradigm that addresses this question. Besides, the data were tested for differences that might be based on the gender of the speakers.

Data from 10 native speakers of German and 10 native speakers of Russian were gathered. Four parameters from the electroglottographic wave form combined with the acoustic signal were

analyzed: i. preceding vowel duration duration, ii. voicing into the closure duration, iii. closure duration and iv. release burst duration.

Preliminarily results indicate that German and Russian show similar tendencies: both languages show the tendency to have a longer Voicing into the closure duration in voiced consonants. There are also differences between underlying voiced and voiceless obstruents in word and phrase final position. Overall, a more direct inspection of the dynamics of the vocal folds lends support to the claim that final devoicing does not produce absolute neutralization, but there are phonetic cues that differentiate the production for underlying voiced and voiceless stops. In addition to that there are indicators which imply that phonetic differences produced by native Russian speakers are even more distinctive than those produced by native speakers of German. The findings are discussed in connection to the consequences for theories of speech production and perception.

References:

Baken, R.J. - Orlikoff, Robert, F. 2000. Clinical Measurement of Speech and Voice. (2nd edition). Delmar: Singular Thomson Learning TM.

Baumann, Monika. 1996. The Production of syllables in connected speech. Nijmegen: Benda drukkers.

Charles-Luce, Jan. 1985. "Word-final devoicing in German: effects of phonetic and sentential contexts", Journal of Phonetics 13: 309-324.

Dinnsen, Daniel, A. - Charles-Luce, Jan. 1984. "Phonological neutralization, phonetic implementation and individual differences", Journal of Phonetics 12; 49-60.

Dmitrieva, Olga - Jongmann, Allard. 2005. Phonological neutralization by native and non-native speakers: the case of Russian final devoicing. [Unpublished paper, University of Kansas - Stanford University.]

Fourcin, Adrian. 1986. "Electrolaryngographic assessment of vocal fold function", Journal of Phonetics 14: 435-442.

Hanson, David G. - Gerratt, Bruce, R. - Berke, Gerald, S. - Karin, Rom, R. 1988. "Glottographic Measures of Vocal Fold Vibration: An Examination of Laryngeal Paralysis", Laryngoscope 98: 541-549.

Jassem, Wiktor - Richter, Lutoslawa. 1989. "Neutralization of voicing in Polish obstruents", Journal of Phonetics 17: 317-325.

Jessen, Michael. 1998. Phonetics and phonology of tense and lax obstruents in German. Amsterdam: Benjamins.

Kanibolotskaia, Sofia. 2008. Aerodynamic Study of Word-Final Devoicing in Russian.

[Unpublished MA. thesis, University of Toronto.]
Kenstowicz, Michael. 1994. Phonology in Generative Grammar. Cambridge MA, Oxford UK: Blackwell Publishing.
Ladefoged, Peter. 2001. Vowels and consonants. An introduction to the sounds of languages, Malden, Mass.: Blackwell Publishing.
Ladefoged, Peter. 2003. Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques. Oxford: Blackwell Publishing.
Orlikoff, Robert, F.. 1991. "Assessment of Dynamics of Vocal Fold Contact From theElectroglottogram: Data From Normal Male Subjects", Journal of Speech and Hearing Research 34: 1066-1072.
Port, Robert, F. - O'Dell, Michael, L.. 1985. "Neutralization of syllable-final voicing in German", Journal of Phonetics 13: 455-471.
Smith, L. Caroline. 1997. „The devoicing of /z/ in American English: effects of local and prosodic context", Journal of Phonetcs 25: 471-500
Warner, Natascha - Jongmann, Allard - Sereno, Joan - Kepms, Rachel. 2004. "Incomplete neutralization and other sub-phonemic durational differences in production and perception: evidence from Dutch", Journal of Phonetics 32: 251-276.

Gradient allophony in Brazilian Portuguese alveolar stops: data from a dialect spoken in Jundiaí-SP

Denise Pozzani de F. Barbosa.
denisepz@gmail.com

The aim of this study is to investigate palatalization of alveolar stops before [i] in a dialect of Brazilian Portuguese, namely, that spoken in Jundiaí, a town in the state of São Paulo. Affricates occur in many dialects of Brazilian Portuguese as allophonic variation in such a context, while in others dialects dental/alveolar stops are undergoing change towards them, as they are considered a prestige variety.

Affricates present some instability in their borders, and a complex temporal structure. Thus, our goal is to describe phonetic detail in the affrication process, in a group of speakers that are implementing such a linguistic change.

Preliminary data showed that affricates have continuous characteristics between alveolar stops and their post-alveolar counterparts. Encouraged by these results, we conducted two case studies with speakers of that dialect.

For each subject, there were three data collection sessions, over a year. The materials consisted of newspaper clips, which were read out aloud by the subjects at a normal speaking rate. Each of the texts contained words with /t,d/ before /i/. In addition, a word repetition task was also recorded. The stimuli were recorded by a phonetician and the following variables were manipulated: word frequency, position of the syllable relative to stress and presence or absence of the affricate pronunciation.

In the analysis, we measure spectral moments (Forrest et al. 1988) and the relative durations of both the frication noise and the vowel. Continuous data on spectral moments can be classified into categories such as alveolar, post-alveolar and intermediate. Nonparametric statistical analysis can be used to evaluate the speaker preferences for one of such ranges. Conservative speakers are expected to prefer the alveolar range, and less conservative ones are expected to prefer either post-alveolar or the intermediate range.

Furthermore, the correlation between the relative duration of vowels and consonants is a good clue to the idea that there is an overlap between consonant and vowel gestures. The possibility of this overlap has been suggested by Albano (1999 and 2001).

Initial data inspection indicates growing presence of affricates in the production of both

subjects from one session to the other. Preliminary spectral moment analysis shows sufficient variation so as to indicate instability in affricate implementation. We hope that statistical analysis of the two data sets sheds light on gradient change in place of articulation.

Gestural Phonology (Browman and Goldstein, 1992, 1995, Goldstein, Byrd & Saltzman, 2006) has proven to be adequate for describing and modeling such gradient processes. Since its theoretical postulates capture the dynamics of speech production over small stretches of time, we hope they also help illuminate the dynamics of language change.

References

Albano, E. C. (1999) O Português Brasileiro e as controvérsias da Fonética atual: pelo aperfeiçoamento da Fonologia Articulatória. In: *D.E.L.T.A.,* São Paulo, v. 15, p. 23-50.

_____. (2001) *O Gesto e suas Bordas*: esboço de Fonolgia Acústico-Articulatória do Português Brasileiro. Campinas: Mercado de Letras; São Paulo: FAPESP.

Browman, C. & Goldstein, L. (1992) Articulatory Phonology: an overview In: *Phonetica*, v. 49, p. 155-180.

_____. (1995) Dynamics and Articulatory Phonology In: R. F. Port & T. van Gelder (eds.). *Mind as Motion*: explorations in the dynamics of cognition. Cambridge, Mass: MIT Press.

Forrest, K.; Weismer, G.; Milenkovic, P.; Dougall, R. N. (1988) Statistical analysis of world-initial voiceless obstruents: preliminary data. In: *Journal of the Acoustical Society of America*, v. 84, n. 1, jul., p. 115-123.

Goldstein, L.; Byrd, D.; Saltzman, E. (2006) The role of vocal tract gestural action units in understanding the evolution of phonology. In: M. Arbirb (ed.) *Action to language via the mirror neuron system*. New York: Cambridge Univ. Press, p. 215-249.

Approaches to the synchrony and diachrony of back fricatives in three dialects of Spanish

Diego Jiquilin-Ramirez
diegojramirez@gmail.com

This ongoing investigation examines the production of back fricatives in Spanish. From a synchronic perspective, the aim is to understand the linguistic changes of these sounds. From examples of back sounds in other Romance languages, we come to Latin, where "aspiration" had two different statuses: i) allophonic variation between a fricative – associated with people of high social status – and its absence, and ii) aspiration after occlusive consonants, a phenomenon borrowed from Greek. Spanish back fricative sounds, however, did not originate from either of these two cases. Historically, there was a glottal fricative phoneme from Latin [f]. Nowadays this is no longer a phoneme, in Jakobson's phonological terms, since there has been a dephonologization that only survives in some dialects. In current Spanish phonological inventory, there is at least a velar fricative (e.g. ca[x]a) - the result of phonologization of the other Latin sounds (such as a palatalization of "l" or groups like "cl", for example). Another back fricative sound is produced in dialectal variety: a glottal fricative (e.g. ca[x]a ~ ca[h]a), often found in Andalusian or Caribbean areas. What caused the appearance of these sounds? Or rather, why a phonologization of the back fricatives took a place in modern Spanish? Today, back fricative sounds are produced at least in three points of the vocal tract, depending on dialect. In order to examine these three varieties, we recorded samples from Madrid, Paraguay and Caribbean speakers. Three subjects represented the Madrid (Spain) dialect, one woman and two men. As representatives of the Paraguayan dialect, we recorded two female informants, both from Asuncion. We have more Caribbean speech samples than others, because this dialectal area embraces many countries: the subjects are four Colombians, three women and one man, all from Bogota; one Venezuelan male, from Caracas; one female from San Miguel, El Salvador; one male informant from San Jose de Costa Rica, Costa Rica; and one Cuban male coming from La Havana. All informants have a high educational level with age varying from 25 to 35 years. As structural independent variable we chose the sounds' vocalic context of occurrence since articulatory activity in posterior areas, besides being influenced by dialectal variety, varies according to its occurrence between [a]_[a],[i]_[i] or [u]_[u]. The target word was introduced in a non-forged carrier sentence, taken from real texts. Words in which the articulation context of the back fricative is [a]_[a] are much more frequent than [i]_[i] production context. This context, in turn, is much more common than [u]_[u]. In order to deal with the difficulty of finding

infrequent vowel pairs, we used a search tool provided by the *Real Academia Española* (RAE), called CORDE (*Corpus Diacrónico del Español*). We chose a diachronic corpus because we could expand the occurrence of words. Therefore, among the target words, a minority, we find some archaic vocabulary. Since the informants are highly educated subjects, there was no difficulty for them in reading the more archaic sentences. Even so, subjects performed readings prior to recording. The corpus was collected at Phonetic Laboratory of the Consejo Superior de Investigación Científica (CSIC), located in Madrid, Spain. By now, we have only a qualitative analysis of the spectrograms, since it is very difficult to find an acoustic-statistical parameter that characterizes alone the productions in posterior areas of the vocal tract. Qualitatively, we speculate that the formant oscillations of such fricatives are a case in point. According to Ohala (1960), formant oscillation, compared to the adjacent vowels, can offer clues about of the articulation place of the consonant. Quantitatively, it is interesting to analyze the fricative duration or its spectral peaks. Following Forrest *et al* (1988) and Jongman *et al* (2000), we preliminary tested the four spectral moments. Results are as yet tentative, but they seem to reflect the vowel context. Can dynamics explain why?

Correlation between the findings on the nasal air emission test and nasopharyngoscopy in patients with operated cleft lip and palate

Fabiana Penido * , Renata Noronha, Kisy Caetano,
Marisa Marisa, Camila Di Ninno and Ana Teresa Britto.
* fabianapenido@gmail.com

Objective: To verify whether the information collected in the nasal air emission test are compatible to the nasopharyngoscopy findings, helping the assessment of the velopharyngeal mechanism functionality.

Methods: The findings of nasal air emission test and nasopharyngoscopy were analyzed in 21 individuals of both genders, over the age of 8 years, operated for post-foramen and transincisive foramen fissure. The correlation was carried out observing whether or not, in the face of a velopharyngeal gap visualized in the nasopharyngoscopy, misting of the Glatzel mirror was noted in the nasal air emission test. Data were statistically analyzed using the Kappa index.

Results: An agreement was observed between the tests ($p<0,05$), and it was considered higher during blowing, emission of isolated stops and fricatives in words and phrases (Kappa: 0,897; 0,887; 0,774 and 0,774, respectively), than for emission of stops in words, isolated fricatives and stops in phrases (Kappa: 0,691; 0,640 and 0,488, respectively). Incompatibility occurred in a few cases, in the presence of a small gap, complete closing and complete closing with bubbling. Regarding the tendency of velopharyngeal closing, it was found a predominance of the coronal type (48%), and improvement was noted in 62% of the subjects with the use of therapeutic test.

Conclusion: The validity of the nasal air emission test when compared with nasopharyngoscopy was observed. It is, however, important to emphasize that the nasal air emission test must be carefully applied and it shouldn't replace the nasopharyngoscopy, which is conclusive for surgical and therapeutic planning.

Keywords: Cleft palate/surgery; Endoscopy/methods; Velopharyngeal insufficiency/diagnosis; Speech therapy

Prosodically-induced variation in acoustic parameters of oral vowels in Tikúna

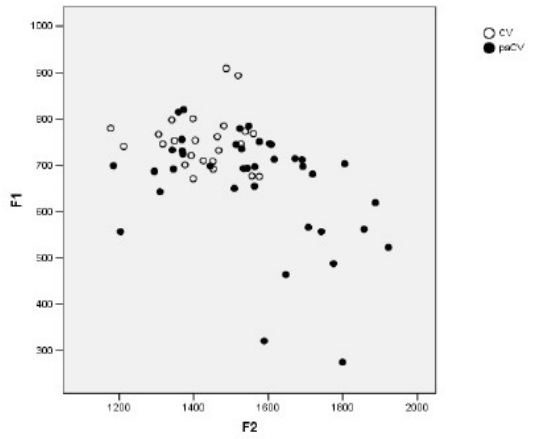
Fernando Carvalho
fernaoorphao@gmail.com

The focus of the present paper is a controlled study on the acoustic properties of the oral vowels of the Tikúna language (isolated; Brazilian, Colombian and Peruvian Amazon) and how these properties change as a function of the prosodic context. As such, our results are compared with those of the pioneering studies on this aspect of the language's sound structure carried by Soares (1984, 1992).

The design of the data elicitation protocol followed a well-known procedure, employing the construction of nonsense words (phonologically possible in the language, but non-existing) which were inserted in invariant carrier-sentences (cf., e.g. Zue 1976; Recasens 1991). Two sets of sentences were constructed: one in which the target syllables occur in stressed position and the other in which they occur in an unstressed one. The invariant carriersentence, meaning "This is a ____" was written in Tikúna orthography and presented in random order to 4 subjects (3 male, 1 female; mean age: 39 years old) all literate native speakers of the language, living in the Ticuna de Santo Antônio reservation, near Benjamin Constant, Amazonas, Brazil. Each sentence, containing the non-sense target word, was read once by each subject and recorded. Some subjects had to read some sentences two or three times until a more natural speech rate was attained.

Formant values, duration and F0 were computed for each token using the relevant estimation tools and the LPC-based algorithm (*Burg*) for formant extraction of the *Praat* software.F0 was double-checked by using the *Voice Report* in order to check glottal period measurements along with the F0 values reported by *Get Pitch.* Optimization procedures to the arguments of the algorithms were applied as a function of speaker sex. The female formant values were normalized by the application of the Nordström-Lindblom normalization procedure (Yang 1996).

Results indicate systematic changes in all the computed parameters as a function of the prosodic condition. Vowel quality changes occurred especially along the dimension defined by F1, as shown in figure 1 below for the tokens of vowel /a/:



In the figure above, black dots indicate the tokens in the unstressed condition, with the white dots indicating the stressed ones. The most striking result was that obtained by the two phonemic posterior vowels /o/ and /u/ whose realizations are largely overlapping in the unstressed condition, while in the anterior portion of the vowel space, no similar pattern obtains between /e/ and /i/. A hypothesis was entertained according to which the /o/ - /u/ contrast is neutralized in non-stressed position. An exhaustive search through an extensive vocabulary of the language (Anderson 1958) confirmed that /o/ does not occur in unstressed position in Tikúna, with /u/ being the exclusive back vowel in this context. The results agree with the results of Soares (1992) concerning the absence of back open vocoids in short syllables in Tikúna. This neutralization process has not been described previously for the language.

Average F0 and average duration were larger in all vowel categories in the stressed as opposed to the unstressed condition, a conclusion in full agreement with the previous descriptions of the language's main stress as implemented by means of increased vowel duration (Soares 1992). The ratios between the average duration in the two conditions showed stressed vowels to be twice as long as unstressed ones. One of the most intriguing conclusions is that the widely-attested patterns of Intrinsic Vowel Pitch and Intrinsic Vowel Duration were observed for all pairs of vowels opposed by height and in all prosodic conditions *except* for the pair /o/-/u/ in the unstressed context, the same context where a superposition in realization and phonological neutralization was observed. This may be evidence to the effect that such patterns of co-variation among phonetic parameters are subject to language-particular phonological conditions.

References

Anderson, Lambert (1958) “Vocabulario breve del idioma Ticuna”. *Tradición*, 8 (21), 53-68.

Recasens, D. (1991) “An Electropalatographic and Acoustic Study of Consonant-to-vowel Coarticulation”. *Journal of Phonetics* 19, 177-192.

Soares, Marília Facó (1984) “Traços Acústicos das Vogais em Tükuna”. *Cadernos de Estudos Lingüísticos* 7, 137-175.

Soares, Marília Facó (1992) O Suprasegmental em Tikúna e a Teoria Fonológica. Doctoral Dissertation, UNICAMP.

Yang, B. (1996) “A comparative study of American English and Korean vowels produced by male and female speakers”. *Journal of Phonetics* 24, 245-261.

Zue, V.W. (1976) Acoustic Characteristics of Stop Consonants: A Controlled Study, Ph.D. Thesis. MIT.

Instrumental analysis of vowel and plosive duration in Brazilian Portuguese

Francisco Meneses and Vera Pacheco.
chico.meneses@gmail.com

This paper aims at assessing emphasis, pause and syntactic position effects on plosive and vowel duration in Brazilian Portuguese (henceforth BP). This work was conducted as an honor's thesis study.

For the analysis, three experiments were conducted: (I) the first was designed to assess the pause effect, (II) the second was intended to analyze the effect of stress, (III) the last experiment was intended to assess the influence of syntactic functions (NP and VP) on the segments duration.

All target words were inserted into carrier phrases. For Experiment I, the words were filled at a distance and in the vicinity of pauses. For the second experiment, speakers were instructed to apply emphasis before, after and on the target word. In Experiment III, the words were inserted in two different sentential positions: NP and VP.

Three subjects were recorded (one man and two women). They are all from Vitoria da Conquista, a city in Bahia, a state of Brazil. A corpus of words with CV and CV.CV structure was created. Onset position was filled with plosives. The vowels / a, i, u / occurred in peak position.

The recordings were made in an acoustically treated environment, in a single session, with five repetitions per subject. Duration measures were obtained from Praat (Boersma & Weenink, 2002).

Statistical analysis was performed with BioEstat (Ayres et al, 2001). An Analysis of Variance (ANOVA) was used for cases where there was no difference in variance. Kruskal-Wallis test was used for cases with unequal variance. Alpha was settled at 0.05.

The results show a strong relationship between pause and segmental duration.

Moderate relationship was found for emphasis. Both contexts, to a greater or lesser extent, display consonant compensatory reduction due to the lengthening of the adjacent vowel (Klatt, 1976). Compensatory reduction is considered a coarticulatory phenomenon (Fowler, 1981).

The results point to the possibility that segmental duration has close links with factors beyond the level of the syllable and the sentence. Duration is strongly influenced by prosodic and conversational factors especially pause.

References:

AYRES, M., AYRES JR, M. AYRES, D. L, & SANTOS, A.LS. Bio Estat 2.0, CNPq, Belém. 2001.

BOERSMA, P. & WEENINK, D. Praat software. Versão 4.0. The Netherlands, Amterdam. 2002.

FOWLER, C. A. A relationship between Coarticulation and Compensatory Shortening. *Phonetica* 38: 35-50. 1981

KLATT, D. Linguistics use segmental duration in English: acoustical and perceptual evidence. *The Journal of the Acoustical Society of America* 59:1208-1221. 1976.

Segmental reduction in sequences of high vowels

Gustavo Augusto Mendonça Almeida.
gustavocook@hotmail.com

This paper aims to examine segmental reduction in sequences of adjacent high vowels in Brazilian Portuguese (BP). Santiago (2005) reports a number of cases where vowel reduction involving sequences of adjacent high vowels occurred. She observed that a high front vowel in a primarily stressed syllable followed by an unstressed high back vowel was reduced so that the vowel in primarily stressed position remained. Examples which illustrate this phenomenon are: a) r[iu] > r[i] *rio* 'river', b) me[iu] > me[i] *meio* 'half' and c) baix[ĩỹ˜u] > baix[ĩ] *baixinho* 'short (diminutive)'. All cases investigated are related to nominal morphology. Cases as illustrated in (a) show that a high front vowel followed by a high back vowel is reduce so that only the high front vowel occurs.

Cases as illustrated in (b) show that a diphthong ending in a front glide which is followed by a high back vowel is reduced so that only the diphthong occurs. Cases as illustrated in (c) show that a sequence of (high front nasal vowel + nasal palatal glide) followed by a high back vowel is reduced so that only a high front nasal vowel occurs.

Interestingly, all the reduced patterns examined by Santiago (2005) have correlate forms in BP here no reduction is reported. For example, for cases such as in (a) the forms which undergo reduction such as r[iu] > r[i] *rio* 'river' may be compared to plain forms such as r[i] *ri* '(s/he) laughs'. Cases in (b) such as me[iu] > me[i] *meio* 'half' might be compared to plain forms such as l[ei] *lei* 'law'. Cases in (c) will be divided in two subgroups. Group (c1) will involve diminutive forms that are very productive in PB baix[ĩỹ˜u] > baix[ĩ] *baixinho* 'short (diminutive)' and group (c2) will involve non-diminutive forms such as viz[ĩỹ˜u] > viz[ĩ] *vizinho* 'neighbour'. The two groups, (c1) and (c2), were examined separately by Santiago (2005) and thus we decided to examine them in an analogous way to her. Both groups (c1) and (c2) may be compared to plain forms such as cap[ĩ] *capim* 'grass'.

Santiago's analysis was based solely on auditory examination. Thus, a closer analysis of her reported cases of vowel reduction deserves further consideration specially with regards to experimental analysis. This paper intends to be a contribution towards this gap in the literature.

The analysis to be presented follows major guidelines of Laboratory Phonology

(Pierrehumbert, Beckman & Ladd 2000) The main acoustic parameter to be analysed is duration. Formant values and intensity will also be considered. Exemplar Models (Johnson 1997, Pierrehumbert 2001) and Usage-Based Phonology (Bybee 2001) will be adopted to address the lexical organization of the data we examined.

Results indicate that duration is a major parameter to distinguish between cases where vowel reduction occurs and the regular cases. Thus, for example, in a case such as r[iu] > r[i] *rio* 'river' which undergoes vowel reduction the high vowel presents a longer duration than cases with a plain form such as r[i] *ri* '(s/he) laughs'. This indicates that fine phonetic detail distinguishes cases that present vowel reduction from cases with plain form. Therefore, it appear that phonological encoding must include phonetic detail. Another interesting result is that diminutive forms, which have morphological structure, differ from non-diminutive forms when undergoing reduction. That means that although a high front nasal vowel occurs in diminutive forms such as baix[ĩỹ̃u] > baix[ĩ] *baixinho* 'short (diminutive)', as well as in non-diminutive forms such as viz[ĩỹ̃u] > viz[ĩ] *vizinho* 'neighbour' and also in plain forms such as cap[i] *capim* 'grass' they present specific properties. This result may indicate that phonological encoding may be related to morphological information (Bybee 2001, Lusiwiecz 1992).

Future research may investigate the role played by type and token frequency in the cases investigated in this paper. A closer examination of the prosodic environment where vowel reduction takes place might also be of interest since the primarily stressed vowel which remains from vowel reduction may occur in the foot domain either as a weak or as a strong constituent.

References

BYBEE, J. 2001. *Phonology and Language Use.* CUP. Cambridge.

JOHNSON, K. 1997. *Speech Perception without speaker normalization*. In: Talker variability in speech processing. K. Johnson & J. Mullenix (eds). 145-65. Academic Press. San Diego.

LOSIEWICZ, B. 1992. The effect of frequency on linguistic morphology. Master's dissertation (manuscript). The University of Texas at Austin. Austin.

PIERREHUMBERT, J. 2001 *Exemplar dynamics: Word frequency, lenition, and contrast*. In J. Bybee and P. Hopper (eds.) Frequency effects and the emergence of lexical structure. 137-157. John Benjamins. Amsterdam.

PIERREHUMBERT, J., M. BECKMAN & R. LADD. 2000. *Conceptual Foundations of Phonology as a Laboratory Science*. Phonological Knowledge - Conceptual and Empirical Issues. N. Burton-Roberts, P. Carr & G. Docherty (eds). 273-303. OUP. Oxford.

SANTIAGO, E. C. 2005. *Alteração segmental em seqüência de vogais altas no português de Belo Horizont*e. Master's dissertation (manuscript). UFMG. Belo Horizonte.

About the linguistic consciousness of prestige and stigmatization of /tʃ/ allophones in first grade children belonging to schools of different socioeconomic groups in Concepción, Chile.

Jaime Crisosto.
jcrisosto@udec.cl

Introduction. In Chile, the allophonic variation of /tʃ/ becomes important in relation with its manner of articulation, more than its place, although this has been object of analysis too, but in lesser extent. The fricativization phenomenon of /tʃ/ obeys to a principle of laxness in the consonantic articulation [1] and it has been recognized in every study about the allophony of this phoneme in Chile, during the second half of the last century. From a social point of view, the variations of this phoneme are enormously outstanding; thus, the most of people who pronounce /tʃ/ only with affricate articulations is very sensitive to the fricatives of their interlocutors and it's not unusual that when facing such phonetic conduct, they react with damning opinions or derisive imitations [2].

Theoretical Framework. It has been described in the literature a multiplicity of widely used allophones of /tʃ/ which varies in the manner of articulation, especially in the duration of the stop and fricative times. Duque and Tassara (1976) detect, through spectrographic analysis, the apparition of four allophones: one 'adhesive' [tʃ] – with a stop time whose duration is the double than the fricative time-, affricate [tʃ] – with a stop and fricative times during the same-, affricate with fricative tendency [tʃ] - with a fricative time whose duration is the double than the stop time- and fricative [ʃ] –without a stop moment- [3]. Later, in 1999, Vivanco sets up the existence of a fifth allophone that he suggested name 'affricate with prolonged close', symbolizing it like [t¬tʃ], which had born in the bosom of the upper-middle and upper class of Santiago, Chile, like a reaction against the stigmatized fricative variant [ʃ] [4].

Supporting on Labov (1972), Trudgill (1972) and Milroy and Milroy (1985), it can be confirmed that the upper sociocultural levels in each community are those that carry out the models of the accepted uses and those that have the kind that tends to be imitated, in other words, they have the prestigious variants of each linguistic use. Is in this way that, in general, in the socially asymmetric linguistic exchanges, the interlocutor who has the stigmatized variant tends to shift his

own use to the prestigious variant, meanwhile the interlocutor who has the prestigious variant keeps his use. These metalinguistic tasks demand a previous development of the linguistic consciousness, who, according to López Morales, tends to be, however, most noticeable in the upper classes [5].

Nevertheless, this phenomenon has not been much studied in children populations whom whit a linguistic consciousness in development, offer an attractive pole for their revise.

Objectives. To observe if the first grade children are permeable or not to the valuation that acquire some phonological features, to determinate if they have acquired the enough linguistic consciousness to discriminate their own use to the other, adjusting their speech in function of the exposure to a prestigious or stigmatized variants of the /tʃ/ phoneme.

Procedure. The sample was composed of 21 speakers who were in first grade. They were distributed in 3 different schools of 3 dissimilar socioeconomic levels. To each student was applied a test which had two main stages; the first 'Spontaneous Generation', where the children were exposed to a sheet with a specific draw in it. Later, they had to tell out loud what was shown to them in the sheet. The draws represented words with the /tʃ/ phoneme in its structure. In the other stage, 'Direct Repetition', the examiner told them a word and they have to repeat it immediately. The words used at this stage were the same that in the first step, but here the variant of the /tʃ/ phoneme changes according to the predominant variant of the child, noticed in the first stage by the examiner, in such a way that here was exposed to the opposite allophone.

The words used -12 in total- were organized according to the position of /tʃ/ on the phonetic context. In the first four words, the phoneme is found after silence, in the following four, it is found after a nasal consonant and in the last four, /tʃ/ is founded in intervowel position.

Results. A. Distinguishing the use of [ʃ] from the total of articulations of the speakers with the prestigious allophone predominance –who were exposed to the fricative variant- on the 'Direct Repetition' stage; in the upper socioeconomic level, in a 5.9% of the total of pronunciations appears the stigmatized allophone, in the middle socioeconomic level, in a 2.7% and in the low socioeconomic level, in a 35.4%.

B. Distinguishing the use of [tʃ] from the total of articulations of the speakers with the stigmatized allophone predominance –who were exposed to the affricate variant- on the 'Direct Repetition' stage; in the middle socioeconomic level, in a 49.9% of the total of pronunciations appears the prestigious allophone, and in the low socioeconomic level, in a 36.9%. There are not data for the upper socioeconomic level, because there are not speakers who use the stigmatized variant predominantly at this level.

Discussion. It can be observed that the shift from the prestigious variant to the stigmatized one is, in percentage terms, minimum, always lower than the 50%. The fact that the distribution was

not completely stratified can be explained because a ‘resilience’ effect of the speakers from the middle socioeconomic level, who, accustomed to hear the fricative variant, reinforce the own. This reinforcement does not occur in the upper socioeconomic level, where the stigmatized variant is always absent. So, the speakers who use predominantly [tʃ] in their speech do not tend to shift their variant to the stigmatized one when they face to a interlocutor who does use it, what it can be explained because they have a ‘linguistic training’ which let them take conscience –linguistic consciousness- about their own use in contrast with the other, keeping their speech because they consider it adequate.

On the other hand, the speakers who use [ʃ] predominantly in their speech do not tend to shift completely their own fricative variant to the prestigious one when they face to a interlocutor who does use it.

Conclusion. Notable differences exist in the use of the allophonic variants of the /tʃ/ phoneme, according to the socioeconomic level. The position that the speakers take on having faced the variant opposite to the one that they use demonstrates a more developed consciousness in the most well-off sectors.

References

[1] Lenz, Rodolfo (1940-1892), “Estudios chilenos. (Fonética del castellano de Chile)”. Amado Alonso y Raimundo Lida (eds.)., El Español en Chile, BDH VI, Bs. As. Instituto de Filología, 85-208.

[2] Valdivieso, Humberto (1998-1999). “La variable fonológica /ĉ/ en Concepción”. BF 37, 1199-1209.

[3] Duque, Carlos y Gilda Tassara (1976). “Análisis acústico de las realizaciones de /ĉ/ en Valparaíso”. Signos 9, 13, 5-31.

[4] Vivanco, Hiram (1998 – 1999). “Análisis fonético acústico de una pronunciación de ‘ch’ en jóvenes del estrato social medio-alto y alto de Santiago de Chile”. BF 37, 1257-1269.

[5] López Morales, Humberto (1992) ”Style, sex and linguistics consciousness”, en Moreno Fernandez ed. 1992, 43-54.

At the intersection of music and language: stylized intonation

Jeremy Day-O'Connell.
jdayocon@knox.edu

Music and language are universal activities of human culture that share a host of basic sonic resources (including pitch, rhythm, and timbre) and perceptual and cognitive faculties (including pattern matching, categorization, syntax and semantics). These and other similarities have often led to the speculation that music and language were closely related in the evolutionary history of our species--two branches of a single, primal form of human communication.[1] Scientists have just begun to study the precise relationship between music and language; indeed, recent work has spawned a veritable discipline unto itself, unifying subfields from musicology, linguistics, and psychology.[2] I wish to make a unique contribution to this emerging discipline through the study of a neglected phenomenon located squarely at the intersection of music and language: "stylized intonation," that is, playful or attention-seeking speech intoned in deliberate, discrete dyads.

Stylized intonation is chiefly used in infant-directed speech ("Yoo-hoo," "Bye-bye") but is regularly recruited in adult-directed speech for endearing (or mock-endearing) effect; it is also used in long-distance calls ("Din-ner!"), taunts ("Nya-nya!"), and *ad hoc* group chants ("Airball!").

Linguists have described this phenomenon variously, as "stylized intonation," the "calling contour," "vocative chanting," and "stereotyped language"[3] and have, moreover, noted its use in several linguistic cultures.[4] It has also been cited (albeit only casually) by many musicologists, who concur with vague claims of universality.[5] Importantly, all these observers inevitably cite a specific musical interval in connection with this linguistic formula: the minor third.

Despite the widespread acknowledgement of stylized intonation and the provocative suggestion of a musical universal (the minor third), however, scholars have never studied this phenomenon systematically but have contented themselves with largely anecdotal evidence.

The lack of serious literature on the topic can perhaps be explained by its precarious location at the intersection of disparate disciplines. The lack is deeply regrettable, however, as careful investigation would likely bear important implications for more fundamental questions in anthropology and evolutionary psychology, such as the existence of similarities among the musical scales of the world (albeit very rough similarities [6]), and indeed, speculations about the origins of

human language and music-making.

My project will contribute to the field of intonational phonology by collecting and analyzing intonational data with attention to objective measurements of pitch. The stability and clarity of pitch in stylized intonation will provide a rich basis for the study of pitch more generally in language. [7] Above all, the project's assessment of stylized intonation cross-culturally pertains to the understanding of universals in both language and music, an understanding that impinges on broader questions concerning the origins of these two cultural activities, and the relationship between them. The results will inform diverse fields, from animal communication, to psychoacoustics, to anthropology and human evolution.

My work thus far has involved two studies in laboratory phonetics: an examination of the "knock-knock joke" in American English (one straightforward example of stylized intonation, which I choose for its ease of elicitation); and a study of target words in a variety of sentencetypes (including stylized interjections). These studies have confirmed the distinctiveness of stylized intonation along intonational, durational, and timbral dimensions and provide a starting point for cross-cultural work. I have also begun to conduct pilot interviews in order to assemble a vocabulary of stylized expressions in a number of languages (including Brazilian Portuguese). Further work will require on-site (or, possibly, online) controlled elicitations of stylized intonation in a great variety of languages, in order to assess and perhaps refine the universalist assumption.

Notes

[1] The debate has spawned two poles, represented on the one hand by Pinker (1997), who holds the view that music is an outgrowth of the (more clearly adaptive) human capacity for language, and on the other hand by Mithen (2006), who believes that music was, as it were, the first language. Other notable speculations and theoretical treatments of the subject include Cross & Morely 2008; Cross 2006; Cross 2003; Jackendoff 2003; Huron 2001; Brown 2000; Richman 2000. A diverse collection of essays on the origin of music is contained in Wallin et al., eds. 2000.

[2] This emerging discipline can be discerned, for instance, in recent international conferences ("Music and Language," Cambridge 2007, and "Music, Language, and the Mind," Boston, 2008) and recent monographs (such as Patel 2007 and Mithen 2006), as well as several books treating language in the course of an examination of music perception and cognition, such as Zatorre & Peretz, eds. 2001; Aiello 1994; Jones & Holleran 1992.

[3] Respectively, Ladd 1978; Ladd 1996 and Abe 1962; Liberman 1979; Chao 1956. Stylized intonation is mentioned in much the same way in all monographs on phonology or tone. Gibbon (1976) gives a particularly careful discussion of the meanings of stylized intonation; as is typical,

however, this fails to contain any discussion of pitch.

[4] Various authors in Hirst and Di Cristo, eds. (1998) mention stylized intonation in English, German, Dutch, French, Romanian, Japanese and Thai.

[5] For instance: Ringer 2001; Alper 1992; Heaton 1992; Brailoiu 1984; Bernstein 1976.

[6] Nettl 2000; Nettl 1956.

[7] As an example, one theoretical question in linguistics that remains somewhat unsettled--from a musical point of view, bafflingly so--is the appropriate scaling of pitch measurements (e.g., logarithmic versus arithmetic). Most phonologists use one or the other scaling, generally without justification and with little regard for the importance of the issue. Cruttenden (1997) at least takes the question head on, but makes an unsatisfying simplification by taking logarithmic and arithmetic measurements as equivalent in the frequency range of human speech. (In fact, the common vocal range of adult speakers is wide enough that the two measures can potentially differ considerably.) Rietveld and Gussenhoven (1985) present evidence to show that *perception* of intonational prominence is better predicted by excursion size as measured arithmetically than logarithmically, but more recent work (Nolan 2003) supports logarithmic scaling. The present study--again, because of stylized intonation's amenability to objective measurement--will contribute important data relevant to this question.

Gestural and acoustic evidence for the syllabic organization of children's early words

Jill Thorson.
jill_thorson@brown.edu

The timing of acoustic segments with articulatory gestures in early word productions is a previously unexplored area of child language acquisition. The driving motivation for this research is to look at how children form early syllabic structures and where they are producing boundaries between segments, looking at both the gestures of the articulators and the acoustics of the speech. As previous work has shown, acoustics can exhibit covert contrasts to distinguish segments (Demuth, Culbertson & Alter, 2006). Also, longitudinal corpus data show that some children employ the use of epenthesis and aspiration in such a way as to simplify the syllabic structure of a word, as found in a preliminary study of Naima of the *Providence Corpus* at age 1;2 to 1;5. The question that then arises is if there are covert movements of the articulators as well that may reveal how the phonology of such segments and syllables is being realized during speech. The current study uses ultrasound and auditory recordings of adults (to establish a baseline) and children age 2;0 to 2;6 in order to obtain a better picture of how early monosyllabic and disyllabic words are being realized as a function of gestural and acoustic timing, post release noise effects, syllabification and coarticulation.

In the study of covert contrasts, previous work shows that children are often found to omit final (coda) consonants in their speech during early language development (e.g. dog > /dʌ/).

Even once codas begin to be produced they sometimes occur with heavy aspiration, an epenthesized final vowel, or a long interval of silence before the release of the coda (cup > /kʌpʰ/, big > /bigə/, big > /bi: […] g/) (Demuth et al., 2006). As an initial step in the current study, an analysis of the perceptual phonetic transcriptions of Naima's CVC[+stop] sequences was conducted. Naima's age range analyzed was from age 1;2 to 1;5, where she is producing such alterations of epenthesis and aspiration most robustly. Results show that aspiration occurs most frequently in final position across the age range, while epenthesis shifts from a final to medial utterance position preference (this could also be a result of the use of more multi-word sequences at age 1;4 and 1;5). Additionally, aspiration most often follows a voiceless coda consonant, while epenthesis does not follow any such pattern and occurs after both voiced and voiceless coda

consonants depending on the age. These results are shown graphically in Figures 1 and 2 below.

These preliminary findings from the perceptual analysis are ambiguous though with respect to the nature of children's early syllabic representations and beg the question of how children are actually syllabifying such sequences. The goal of this project is therefore to look at the acoustic and gestural underpinnings of children's early CVC words to determine if the final Cs are syllabified as codas (CVC) or onsets (CV.CV). To do this I will carry out 1) an acoustic corpus study of spontaneous productions from the child Naima of the *Providence Corpus*, who showed extensive use of the above processes between the ages of 1;2 and 1;5, and 2) an ultrasound study with five 2-year-olds and five adults looking at the gestural organization of nonce words.

First, an acoustic analysis of the tokens from the preliminary perceptual analysis of Naima is being performed. This part of the analysis examines Naima's aspirated and epenthesized coda consonant productions and their timing relative to the first part of the sequence of the target word (the CV onset/nucleus versus the coda C). Acoustic landmarks are coded in Praat, with specific interest in 1) the duration of the post release noise (the epenthesis or aspiration), 2) the pre-aspiration (or noise at the end of the vowel, 3) the duration of the vowel, and 4) the duration of voicing (Boersma & Weenink, 2009). The purpose of the first part of this analysis is to observe the acoustic realizations of these epenthesized and aspirated productions, as well as the syllabification process.

The second element of this study is the ultrasound experiment that will record both articulatory gestures and speech acoustics during the productions of five nonce word stimuli by five children age 2;0 to 2;6 and five adults (as a baseline for comparison). Subjects are recorded while holding the ultrasound probe under the chin and producing five nonce words elicited during an interactive computer-based word game. The nonce stimuli were chosen based on syllabification and ultrasound preferences (*Stimuli*: /bɑ/, /bɑk/, /bɑg/, /bɑkə/, /bɑgə/). The experiment is still underway with three children already recorded. Future gestural analysis will consist of tracking the tongue shape and height from the vowel to the /k/ to observe any covert contrasts when the /k/ is produced as a coda versus an onset.

Additionally, the stimuli provide a setting to observe how children at this age period are syllabifying such sequences and if they are producing any post release noise similar to that of Naima. The findings will provide much needed information regarding the nature of children's early syllabic organization, and will contribute to building a model of child speech planning and production.

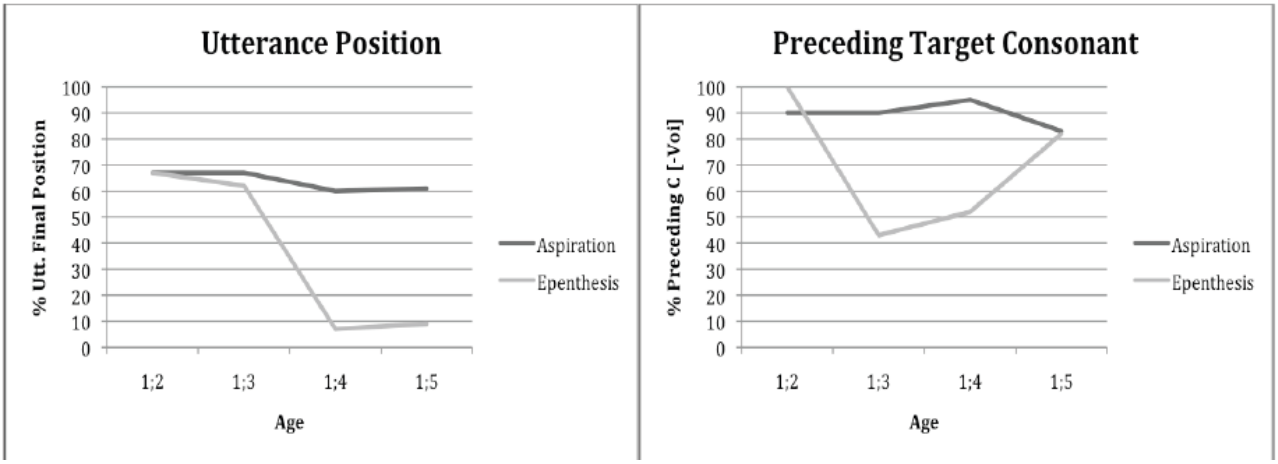


Figure 1 (left): Shows the percentage of tokens that are aspirated and epenthesized as a function of utterance position. The *y*-axis is the percentage of tokens of that category that were produced in final position (versus medial position).

Figure 2 (right): Line graph showing the aspirated and epenthesized tokens as a function of the preceding target consonant. The *y*-axis is the percentage of final coda consonants that had the feature of minus voice (CVC[-Voi]).

*The *x*-axis in both graphs is the Naima's age from 1;2 to 1;5.

References

Boersma, P., and Weenink, D. (2009). *Praat: doing phonetics by computer* (Version 5.1.15) [Computer Program]. URL: http://www.praat.org/ in September, 2009.

Demuth, K., Culbertson, J., & Alter, J. 2006. Word-minimality, epenthesis, and coda licensing in the acquisition of English. Language & Speech, 49, 137-174.

Procedures for the Analysis of Child Vowels and Obstruents in Brazilian Portuguese

Larissa Rinaldi
rinaldi.larissa@gmail.com
Lafape – IEL – Unicamp

This master´s thesis study aims at creating a reference database for the study of vowels and obstruents in child speakers of Brazilian Portuguese (henceforth BP). The subjects are between 5 and 7 years old, with no history of speech disorders. To this end, we have developed a special tool for collecting the subjects´ speech.

The methodology is centered on a child story entitled "Fairy Tales Go Crazy", and uses a board game to prompt the utterance of 57 words created for the instrument. The total number of target words is 36. They were created by matching each PB obstruent (fricatives and plosives) with the vowels [a], [i] and [u]. The methodology allows for collection of all the sounds of PB in initial and medial onset position. This study used the tool to observe obstruents and vowels only in initial stressed position.

Data were collected from 9 children (5 girls and 4 boys) in the final stages of language acquisition, who were selected as randomly as possible. The researcher asked each child to say the target words embedded in a carrier sentence. For each child, only one data collection session was enough to record all the data.

In a preliminary analysis, it was noted that the classical acoustic phonetic parameters were not always sufficient to describe the variation and dynamics of the speech signal in child language. Therefore, we have used two types of analysis: a quantitative and qualitative one.

The quantitative analysis made the following measurements: for vowels, relative and absolute duration and the first three formants; for fricatives, relative and absolute duration and the four spectral moments (Jongman et al., 2000); for plosives, relative and absolute duration, voice onset time (VOT) and the burst spectral moments (Forrest et al., 1988). The fricative spectral moments were measured at their center in a 40 millisecond window.

Statistical analysis was performed by a General Linear Model (GLM) with a Repeated Measures Analysis of Variance. Alpha was set at 0.05. In the case of vowels, the dependent variables were the formants and the independent ones were the vowels. As expected, F1 was

effective in differentiating the vowel [a] from [i] and [u], F2 was effective for differentiating the three vowels, F3 was effective only to differentiate the vowel [i] from [u] and [a]. The statistical analysis of fricatives and plosives is still in progress.

The qualitative analysis is a description of the unusual phenomena observed in the spectrograms. It showed that, often, normal children´s strategies are similar to those with language pathologies described by Berti (2006), Freitas (2007) and Rodrigues (2007). The difference between the two groups is the effect caused on the listener. In spite of deviating from adults in acoustic trajectories as seen on spectrograms, normal children seem to produce something which sounds close to what is expected by adults.

Although the analysis of fricatives and plosives is still in progress, a qualitative analysis of vowel formants in spectrograms showed some differences between adults' and children's productions. Vowels formant stabilization for the same word was different in children and adults. Adults' vowels stabilize quickly and their formants are more visible, more defined. Children's transitions seem to be slower than adult's. It seems that the children end up reaching the articulatory target, but in their own timing. This fact can be illustrated by the figures below, where the cursor indicates the vowel stabilization point.

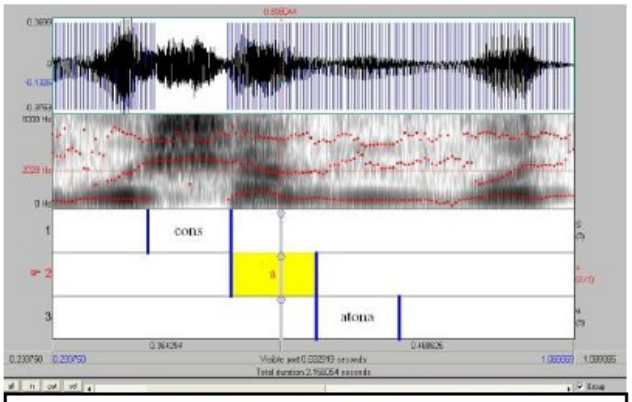


Figure 1: Waveform and spectrogram of the word [Savi], produced by children I., where cons = [S], a = [a] and atona= unstressed syllable = [vi] with cursor at the point of stabilization for better viewing.

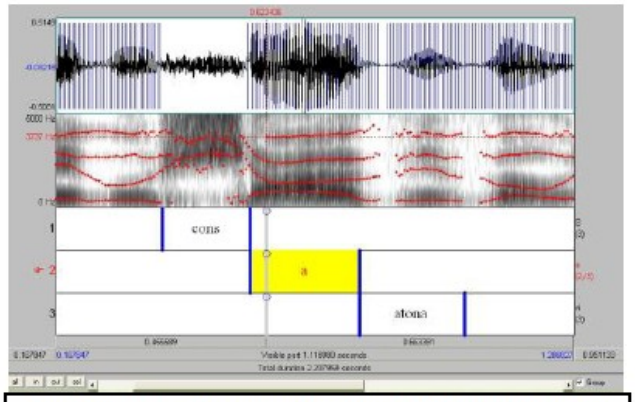


Figure 2: Waveform and spectrogram of the word [Savi], produced by adult L., where cons = [S], a = [a] and atona= unstressed syllable = [vi] with cursor at the point of stabilization for better viewing.

This phenomena may be illuminated by Gestural Phonology (Browman & Goldstein, 1992; Ball & Kent, 1997; Scobbie, 1998; Albano, 2001; Kent & Read , 2002; Shadle, 2006; Goldstein, Byrd & Saltzman 2006). We hope this approach may shed further light on this work.

Bibliography

Albano, E.C., 2001. O gesto e suas Bordas: esboço de fonología acústico-articulatória do português brasileiro. Campinas, SP. *Editora Mercado de letras*

Ball, M.J. & Kent, R.D., 1997. The New phonologies. Developments in clinical Linguistics. Wisconsin, *Singular publishing Ltd*.

Berti, L., 2006. Aquisição incompleta do contraste entre /s/ e /S/ em crianças falantes do português brasileiro. Tese de doutorado inédita, Lafape, IEL, Unicamp

Browman C.P., Goldstein L., 1992. Articulatory phonology: an overview. *Haskins Laboratories, New Haven.*

Freitas, M.C., 2007. Aquisição das obstruintes por crianças de 5-7 anos com queixa fonoaudiológica. Dissertação de mestrado

Forrest, Karen, Gary Weismer, Paul Milenkovic, and Ronald Dougall. 1988. Statistical analysis of word-initial voiceless obstruents: Preliminary data. *Journal of the Acoustical Society of America* 84.115-23.

Goldstein, Louis, Dani Byrd, and Elliot Saltzman (in press). The role of vocal tract gestural action units in understanding the evolution of phonology. (2006) In Michael Arbib, (ed.) *From Action to Language: The Mirror Neuron System*. Cambridge: Cambridge University Press.

Jongman, Allard, Ratree Wayland, and Serena Wong. 2000. Acoustic Characteristics of English fricatives. *Journal of the Acoustical Society of America* 108.1252-63.

Kent, R.D. e Read, C., 2002. The acoustic Analysis of Speech. Iowa. *Singular Publishing Company*

Rodrigues, L.L., 2007. Aquisição dos róticos por crianças de 5-7 anos com queixa fonoaudiológica. Dissertação de mestrado

Shadle, C., 2006. Phonetic, Acoustic. Haskins Laboratories, New haven, Connecticut, EUA.

Scobbie, J.M., 1998. Interactions between the acquisition of phonetics and phonology. Queen Margaret University College, Edinburgh.

Vowel-to-vowel perseveratory dissimilation effect in Brazilian Portuguese mid vowels

Laudino Roces
laudino@gmail.com
Lafape – IEL – Unicamp

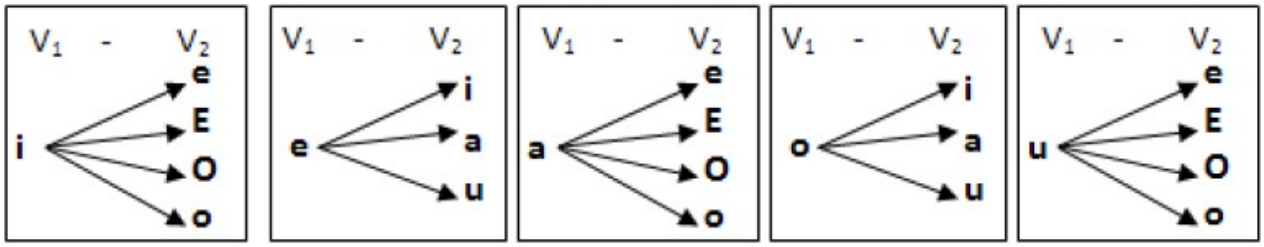
This work is part of a recently finished doctoral dissertation concerning vowel-to-vowel relations in V'CV sequences in Brazilian Portuguese (henceforth BP).

Classical works on vowel-to-vowel coarticulation (Öhman, 1966, Bell-Berti & Harris, 1976, Fowler, 1980) only report assimilatory or neutral effects from neighboring vowels in V_1CV_2 sequences. For instance, the second formant increases if the vowel is adjacent to [i] and decreases if it is adjacent to [u].

This study reports dissimilation effects in V-V relations: formant values increase or decrease contrary to expectation based on adjacent vowel formant configuration. Thus, the second formant decreases if the vowel is adjacent to [i] and increases if it is adjacent to [a].

Brazilian Portuguese vowels /i, e, E, a, O, o, u/ (sampa notation) and consonants /p, t, k/ were used in a disyllabic no-word $pV_1'CV_2$ structure embedded in a carrier phrase "Não, um ____ lindo passou" (No, a beautiful____passed).

V_1-V_2 pairs formed:



Ten female and five male talkers were recorded. Recordings were digitized with a 22 KHz sampling frequency. *Praat* was used to measure F1, F2 and F3.

Variance Components method was used in statistical analyses. /i, a, u/ vowels were used as fixed factors. /e, E, o, O/ were used as dependent variables. Intermediary consonant was used as

covariate. Subject's differences were used as the random factor. The alpha value was 0,05.

Assimilation (AS.), dissimilation (Dis.) effects or neutral (Nt.) were observed.

Stressed vowels´ influences on prestressed vowels (V_2 ⇨ V_1):

Variance Components results - stressed position /i,a,u/ vowels´ influences in prestressed position /e, o/ vowels				
	/e/ - pret.		/o/ - pret.	
	F	*p*	*F*	*p*
F1	9,408	0,001	8,790	0,001
F2	25,740	0,000	0,145	0,866
F3	28,023	0,000	0,228	0,798

Effect results to vowels in prestressed position						
Vowel	/e/			/o/		
Formant	F1	F2	F3	F1	F2	F3
Significance	S	S	S	S	NS	NS
Results	As.	As.	As.	As.	Nt.	Nt.

Prestressed vowels´ influences on stressed vowels: (V_1 ⇨ V_2):

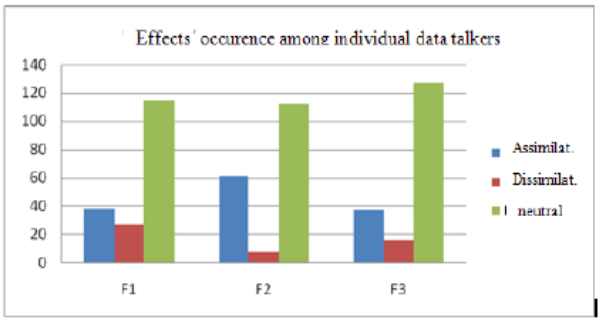
	Variance Components results - prestressed position /i, a, u / vowels´ influences in stressed /e, E, O, o/ vowels							
	/e/ pret		/E/ - ton		/O/ - ton		/o/ - ton	
	F	*p*	*F*	*p*	*F*	*p*	*F*	*p*
F1	3,2962	0,051	2,684	0,086	3,725	0,037	8,790	0,001
F2	5,28664	0,011	7,2611	0,003	1,269	0,297	0,145	0,866
F3	9,73972	0,000	10,779	0,000	0,104	0,902	0,228	0,798

Effect results to vowels in stressed position												
Vowel	/e/			/E/			/ O /			/o/		
Formant	F1	F2	F3	F1	F2	F3	F1	F2	F3	F1	F2	F3
Significance	NS	S	S	NS	S	S	S	NS	NS	S	NS	S
Effect	Nt.	Dis.	Dis.	Nt.	Dis.	Dis.	Dis.	Nt.	Nt.	Dis.	Nt.	Nt.

All mid vowels received significant influences of neighboring vowels (V_1 or V_2). Only anticipatory assimilation effects (As.) occurred in prestressed position. Dissimilation only occurred in stressed position.

Stressed mid vowels /e, E, o, O/ present larger F2 means when the adjacent vowel was /a/ and not /i/.

GLM Anova on individual subject data pointed to a lower frequence of dissimilation, as compared to assimilation or no effect, among the subject.



The existence of dissimilation in V-V relations points to a complex coproduction V-V strategy. This is possibly used to maintain or restore the quality of the vowel in stressed position. Variability in V-V effects shows that V-V relations in BP mid vowels are not entirely predictable.

The complexity observed in BP V-V relations raises many questions to dynamic models of speech production.

References

ÖHMAN, S. (1966). *Coarticulation in CVC utterances: spectrographic measurements,* Journal of the Acoustical Society of America, 39, 151 – 168.

FOWLER, C. A. (1980). *Coarticulation and theories of extrinsic timing control*. Journal of Phonetics, 8, 113 – 133.

BELL-BERTI, F. & HARRIS, K. S. (1976). *Some aspects of coarticulation*. Haskins Laboratories Status Report on Speech Research, SR-45/46, 197–204.

Aspiration of /s/ in Spanish. An account of morpheme boundary processes

Leandro Bolanos.
leandro.bolanos@yale.edu

Spanish is a widely spoken language with many varieties and dialects. The current study focuses on one aspect of Spanish phonology, namely the process known variably as aspiration or debuccalization. Aspiration is a phenomenon whereby the segment /s/ is realized as [h], [ʔ], or even reduces further in different environments. This phenomenon occurs in several varieties of Spanish including the focus of the study, a variety spoken in Central America, specifically Nicaraguan Spanish. The phenomenon occurs in instances of /s/ in final position, be it syllable or other prosodic domain final position (Harris 1983, Hualde 1989, Colina 2006, Wiltshire 2006). Differences arise in the realization of /s/. When followed by a voiceless obstruent or a liquid, /s/ is realized as [h]. When followed by a voiced stop, /s/ is also debuccalized, but the spread glottis is also assimilated to the state of the following stop, i.e. the [h] is also lost. Morpheme boundaries also exhibit interesting interactions. When /s/ occurs in utterance-final position, /s/ variably reduces to [h]. When word-final and followed by a consonant initial word, the same pattern is observed as described above, [h] before voiceless stops and liquids, complete deletion before voiced stops. However, when word-final and followed by vowel initial words, /s/ also reduces. In these cases variation arises, sometimes [h] arises, and sometimes glottal stop. The examples below illustrate the environments described.

Description	Example:		Gloss:
word-internal pre-consonantal:	/sC/	este → ehte	*this*
word-final pre-consonantal:	/s#C/	mas karo → mah karo	*more expensive*
word-final pre-pausal:	/s##/	mas → mah	*more*
word-final pre- vowel:	/s#ˊV/	mas alto → mah alto → maʔ alto	*taller (more tall)*

Such phenomena illustrate several problems. First, the conditioning environments for /s/ aspiration seem to be syllable final. The final /–s#V/ examples constitute testing grounds for phonological theories of syllabification, resyllabification, rule ordering and Stratal Optimality Theory, since across word boundaries, Spanish has been claimed to exhibit resyllabification, as in the data below (Harris 1983, Colina 2006, Wiltshire 2006). The first example shows resyllabification of the liquid [l] across a word boundary. The second example reproduces the aspirated /s/ example above and indicates an analogous description showing [h] now in the onset.

el + oso → [e.lo.so] *the bear*
mas + alto → [ma.hal.to] *taller*

If this is the case, then phonological opacity arises, /s/ aspirates while now analyzed as in the onset. Variants of Optimality Theory that allow for serial derivations provide analyses to such data. The data are interesting also in that words and prefixes seem to behave in similar ways. Prefixes with final /s/ also exhibit aspiration and deletion of /s/ in the same melodic environments as word final and word internal /s/. However, stem final /s/ does not behave in the same way. When a suffix appears on a stem final /s/, the /s/ does not reduce when followed by a vowel. The examples below illustrate these facts.

prefix-s:	de**s**teʧar	[{de**h**}teʧar]	*de-roof*
	de**s**aser	[{de**h**}aser]	*un-do*
stem-s:	{me**s**}$_{st}$+es	[me**s**eh]	*months*
word-s:	tos	[to**h**]	*cough*
	mes	[me**h**]	*month*

In these examples we see clearly that prefix final /s/ as in desteʧar, and word final /s/ as in *tos* behave in similar ways, they both aspirate. While stem final /s/ does not aspirate in all of the same environments. Comparing ‘months’ with ‘cough’ and ‘month’, we see that the plural ‘months’ does not have aspiration of stem final /s/, now that a vowel initial prefix has been realized. This is different than what we see in prefix final /s/, which does aspirate before vowels.

The purpose of the study is to shed light on this data, which has not so far been as completely categorized. Aspiration is a well-studied phenomenon, but the extent and variable realizations of the phenomenon have not been studied in great detail. We attempt to explain the

phenomenon and the effects of different prosodic boundaries with regards to /s/ aspiration and show that Stratal Optimality Theory, that is, Optimality Theory with serial progressions is necessary to account for the data. We also provide analyses of other dialects which show less categorical aspiration, and yet others which show no aspiration and attempt to provide a typology of the possible ways that final /s/ can manifest. We also suggest future work in a framework of Gestural analysis or Articulatory Phonology (Browman & Goldstein 1986, 1989, 1992) to further explain this phenomenon and provide further explanation of the process of /s/ aspiration and its different realizations in different environments.

Applying Recurrence Analysis to Speech

Leonardo Lancia and Susanne Fuchs.
Leonardo.Lancia@univ-provence.fr

1 Purpose

The current study aims to apply to speech production recurrence quantification analysis, a technique designed to detect the presence of recurrent spatiotemporal patterns in multivariate time series. Speech production is indeed a complex dynamical system which is challenging to investigate - at least due to its non stationarity, its physical properties (e.g. soft body motion) and its biomechanics. We think that such an application is particularly fruitful, since recurrence analysis can deal with non stationary time series and it assumes the presence of underlying dynamics (Webber and Zbliut, 1994). We propose a newly developed algorithm which quantifies the acoustic similarity between successive speech signals by comparing the evolution over time of the signal's acoustic attributes, rather than comparing in a pair wise fashion time aligned values. Our work is thus relevant for approaches to speech which assign a key role to temporally changing characteristics of speech sounds.

2 Recurrence analysis

In outline, a m-dimensional process which unfolds in time is represented as a time ordered sequence of n coordinate vectors (V_i, with i={1, ..., n }). The basic step in recurrence analysis is to measure the distance between all the N2 possible pairs of vectors V_i and V_j and to store it in the matrix Di_j. A recurrence plot (henceforth RP, see figure 1a) is a graphic visualization of the obtained matrix, where each distance smaller than a given threshold is plotted as a black dot on an otherwise white graph. Figure 1a shows a RP (top) and the evolution of the periodic bi-variate time-series (bottom) from which it is derived. The main diagonal is called the line of identity or LOI, because dots on this line are obtained comparing each vector to itself.

Since in a periodic process the distance in time between repeated values is constant, a periodic process which is repeated once produces on each side of the LOI a diagonal in parallel to the LOI which is called line of synchrony (LOS). Given the importance of diagonal lines, most measures derived from RPs are statistics computed on the lengths of the diagonal sequences of continuously occurring black dots (i.e. sequences of black dots uninterrupted by white dots).

2.1 Application to speech signals

In order to test the methods on controlled stimuli, we first used synthetic CVCV-speech signals. Secondly, the method was applied to real speech data from a reiterant speech production experiment.

1.) Pairs of synthetic CVCV-speech stimuli were produced in which the two stimuli differ with respect to the order of the consonants or the intensity of one of the vowels. Moreover, we also compressed non-uniformly the duration of one of the utterances or added noise to the whole acoustic signal. Several recurrence analyses were carried out and applied to the pairs of stimuli. Results provide evidence that recurrence analysis is sensitive to all kinds of manipulations, and it is robust against non uniform time-compression and noise.

Figure 1b shows an example of synthetic /pata/ utterances concatenated in a sequence. From the 2 synthetic utterances, 13 mel cepstral coefficients and their first and second derivatives were calculated and served as an input for the recurrence analysis. The RP differs from the one in figure 1a with respect to several characteristics; we are interested in the presence of the gap in the final portion of the LOI. It turns out that this feature is due to manipulation of the intensity of the last vowel of the second utterance.

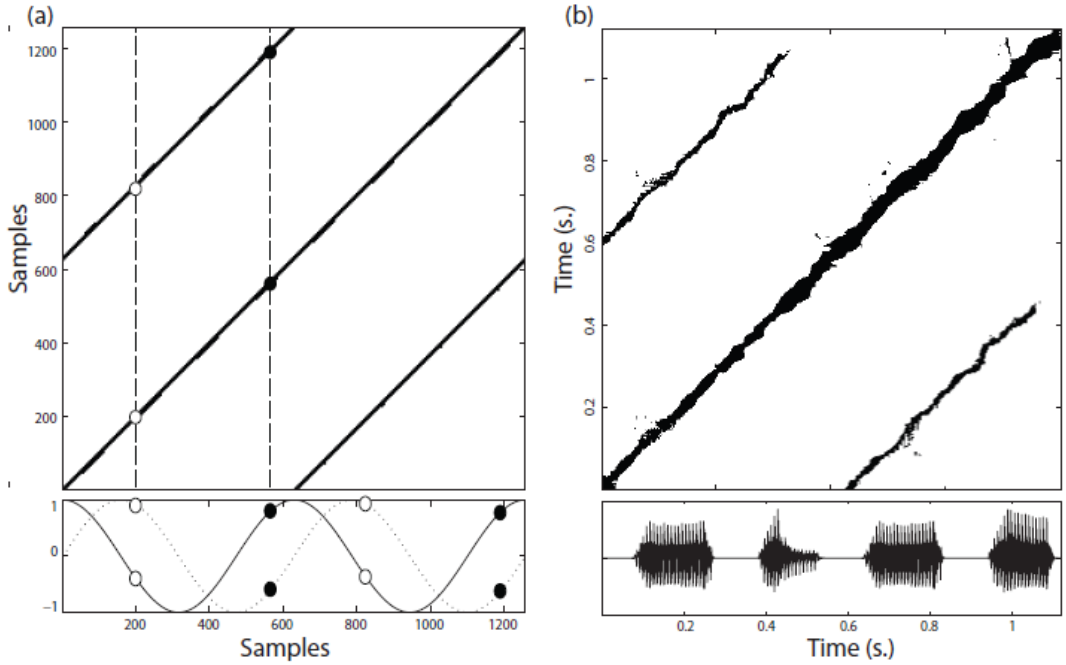


Figure 1a: RP (top) of the two time-series sin(theta) and cos(theta) calculated as theta goes trough the [0; 2\pi] interval twice. Figure b: RP (top) computed onto the trajectories of the acoustic features extracted from the waveform corresponding to a /pata/ utterance repeated twice.

2.) The method was applied to acoustic data recorded during an experiment conducted to study verbal transformation effects. Eight German speakers were recorded by EMA and acoustics. Speakers were asked to produce reiterate /pata/ or /tapa/ sequences with increasing and decreasing speed, guided by a visual metronome. Each pair of consecutive utterances belonging to a sequence has been manually labeled according to the features which change from the first to the second utterance. So far our analysis was carried out for 3 speakers. The behavior of the recurrence analysis shows to be consistent with the manual labeling, and recurrent patterns decrease when changes are observed from an utterance to the following one. Thus, recurrence analysis works for synthetic stimuli and real speech data. In a next step we will apply this method to articulatory data on verbal transformation effects and speech errors.

References

Webber Jr, C., & Zbilut, J. (1994). Dynamical assessment of physiological systems and states using recurrence plot strategies. *Journal of Applied Physiology*, 76(2), 965.

Authentic and acted responses to physical and emotional distress

Lisa Roberts.
lsr501@york.ac.uk

The analysis of speech and sound has an increasing presence within criminal investigations. Given the widespread availability and use of mobile telephones, violent attacks are frequently audio-recorded by victims, witnesses and even the perpetrator(s) themselves.

The development of Forensic Speech Science (FSS), i.e. the practical application of speech and sound analysis within criminal investigations, integrating the methods and theoretical frameworks of phonetics, phonology, dialectology and sociolinguistics, has led to research exploring speaker–specific features such as the overall distribution of vocal features in a given population, and the demarcation of boundaries of the individual's vocal repertoire.

The forensic speech scientist may also be asked to make assessments of the psychological states of victims in recorded attacks. However, such evaluations are currently not advocated by the International Association of Forensic Phonetics and Acoustics (IAFPA), most likely due to a lack of sufficient research on which to form an opinion (IAFPA Code of Practice, clause 9).

This study represents a first step towards research that might ultimately be used to substantiate expert opinions concerning the vocal capabilities of an individual and their emotional state by exploring vocal responses in authentic and acted forensic situations where severe physical/emotional (di)stress is endured. Previous studies investigating emotional speech have concerned speech technology applications (Erickson, 2005) and/or observed vocal cues of emotion in everyday speech (Scherer, 2003). Few attempt to characterise extreme emotion and fewer use authentic data.

The aims are:

1. to investigate the limits of the individual's vocal performance;

2. to identify vocal cues of authentic distress;

3. to compare and contrast authentic and simulated distress.

The data analysed are taken from (i) telephone recordings involving an individual who has been – or is being - subject to a violent attack, and (ii) comparable simulated material involving actors.

The material is subject to both acoustic and auditory-phonetic analyses. Parameters under investigation are chosen following observations from studies using recordings of psychologically stressed aviation personnel (William and Stevens, 1969; Kuroda et al., 1976), and include fundamental frequency (F0), tempo, intensity as well as vowel quality and intelligibility.

Preliminary findings show:

- an increase in F0 overall;
- an increase in variability of F0;
- fluctuations in tempo;
- non-predictable variations in intensity;
- a reduction of vowel quality and intelligibility.

A model is proposed charting changes to these vocal parameters relative to timing of attack. Observations are discussed in terms of their relevance to forensic speech practitioners and further areas of the research are considered.

References:

Erickson, D. (2005) Expressive Speech: Production, perception and application to speech synthesis. *Acoust, Sci & Tech* 4. 317-325.

International Association of Forensic Phonetics and Acoustics. *Code of Practice – section 9*. Accessed March 30 2009 at <http://www.iafpa.net/code.htm>.

Kuroda, I., Fujiwara, O., Okamura, N., and Utsukli, N. (1976). Method for determining pilot stress through analysis of voice communication. *Aviat, Space Environ, Med* 47. 528-533.

Roberts, L. S. (2008) *The Phonetics of Distress*. Unpublished MSc Dissertation, University of York.

Scherer, K.R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication 40* 227-256.

Williams, C.E, and Stevens, K.N, (1969). On determining the emotional states of pilots during flight: An exploratory study. *Aerosp. Med., 40,* 1369–1372.

High vowels in Brazilian Portuguese and British English

Luci Kikuchi
lkikuchi2005@yahoo.com

This research analyses the high vowels /i,u/ in Brazilian Portuguese and British English (RP). These sounds have been chosen due to their high occurrence in natural languages and also because they are part of the three-vowel system: [i,a,u]. The two high vowels represent the maximal contrast, tending to occur in all languages (Roca & Johnson, 1999).

The study aimed to show similarities and differences among the high vowels /i, u/ in the sound systems of two languages: Brazilian Portuguese (BP) and British English (RP). Preliminary studies such as Abercrombie (1967), Jones (1969), Gimson (1970), Ladefoged (2006), Bisol (1989), Cagliari (1998) and Cristófaro-Silva (1999) provided with the preceding descriptive information. These works were relevant in describing articulatory and acoustic characteristics of high vowels and also to describe their phonological or distributional patterns, thus contributing for the phonetic and phonological study of high vowels in general.

Portuguese data showed that the tense vowels /i, u/ have phonological status in Portuguese. That is, in Portuguese we observe phonological contrast between the front and back tense vowels in primarily stressed position as in the words “pira” [pia] and “pura” [pua]. In unstressed position the vowel distribution varies, indicating that stress is a major determinant of vowel properties and distribution in BP. The analysis of formant values will support this claim. In sum, it is posited that in BP primarily stressed vowels qualitatively and quantitatively differ from unstressed - pretonic or postonic - vowels in BP once stress contributes to shaping BP phonology. This result would indicate that Exemplar Models have an important role to play in shaping representations (Pierrehumbert, 2001). This follows from the fact that fine phonetic detail plays a role in determining vowel distribution.

Regarding high vowels in English their occurrence is mainly determined by syllable patterns rather than stress. That is, the tense vowels [i,u] may occur in open syllables as in [si] ‘sea’ or [ju] ‘you’ as well as closed syllables as in [sit] seat’ and [ju] ‘use’. On the other hand, lax vowels in English must appear in a closed syllable: [st] sit’ and [bk] ‘book’. Thus, when formant values are considered tense and lax vowels present distinct distribution. However, there are specific cases in English where a high vowel may occur in unstressed position as in [beli] belly’ or [ju] ‘argue’.

In these cases the formant values of high vowels present a greater variability than in the case of tense and lax vowels. In fact formant values in this later case may resemble values expected for either tense or lax vowels. We suggest that this result provides evidence for fuzzy phonological categories under specific circumstances. In the case analysed in this paper we suggest that the loss of contrast between tense and high vowels in final unstressed position allows for this fuzzy category to appear.

This study has interesting implications for second language acquisition. An investigation of Brazilians learning English and English speakers learning Portuguese may shed some light on how new phonological categories emerges. This issue will be pursued in the future.

Phonological relations between speech and writing in preschool children

Luciana Lessa Rodrigues.
luc.lessa@gmail.com
Lafape – IEL – Unicamp

The aim of this doctoral dissertation study is to investigate links between phonological aspects involved in the acquisition processes of speech and writing in preschool children. We investigate whether there is greater or lesser dependence between writing and phonological/phonetic knowledge.

Writing samples and speech recordings of 12 children (5 and 6 years old) were collected. Three children presented phonological disorders, while the other nine had no history of language disorders. The writing samples were made in the classroom with the teacher. Speech recordings were made in an acoustically treated environment. The researcher asked the children to report on the writing activity. The data were collected monthly for nine months. For this presentation, we have selected June and November data collection sessions.

Our intention was to observe: (a) phonological knowledge underlying the acquisition of writing; (b) whether children with phonological disorders resemble normal children in the phonological hypotheses with which they approach writing; (c) whether there were changes in this respect from June to November.

In the analysis, the occurrence of omissions and substitutions in the writing samples were analyzed according to syllable position (onset, peak or coda). The percent of omissions and substitutions was calculated by dividing the number of occurrences by the total number of possibilities of each position. So far, we have only resorted to speech recordings in order to seek possible explanations for the omissions and substitutions detected in writing. However, we plan to make an acoustic analysis of such data.

The June data yielded the following results: stressed onset - 5% omissions and 2,5% substitutions; unstressed onset - 3,67% omissions and 2,75% substitutions; stressed peak - 9,09% omissions and 0% substitutions; unstressed nucleus – 11,02% omissions and 3,15% substitutions; stressed coda – 14,81% omissions and 0% substitutions; unstressed coda – 19,23% omissions and 15,38 % substitutions. It is clear that both stress and syllable position affect children's writing

errors.

The November data yielded the following results: stressed onset – 1,94% omissions and 1,94% substitutions; unstressed onset – 5,52% omissions and 6,13% substitutions; stressed peak – 5,15% omissions and 0,74% substitutions; unstressed peak – 4,29% omissions and 4,91% substitutions; stressed coda – 21,21% omissions and 10,61% substitutions; unstressed coda – 17,39% omissions and 13,04% substitutions.

Thus, little improvement was observed. In both data collection sessions, there was a higher occurrence of omissions and substitutions in coda. Long words and phrases are responsible for most onset and peak omissions, as in "desatamento" for "desmatamento" and "camada de ozono" for "camada de ozônio". Substitutions in onset were related to orthographic ambiguity.

In November, 3 subjects, including 2 with phonological disorders, were responsible for most of the omissions. It should be recalled that, in June, we had not observed more difficulties in children with phonological disorders as compared with normal children, but, in November, the bulk of the omission errors was produced by them.

Substitutions remained about the same for all subjects in both sessions. It should be recalled that such errors were mainly related to orthography.

Through the results and discussion presented, it can be said that the acquisition of speech and writing are intertwined. In many instances, phonetic detail seems to explain phenomena detected in writing. However, other factors, which have emerged in a post-hoc analysis, are currently under examination.

Elettoglotographic study from Brazilian Portuguese fricative voiced sounds

Luciana Pagan-Neves and Haydée Wertzner.
lucianapagan@usp.br

Introduction: EGG is a non-invasive method executed by the positioning of two electrodes one on each side of the thyroid cartilage. Electric conductivity is measured at the vocal folds level during phonation. Changes in opening and closing movement from the vocal folds are registered in a chart represented by a change in the amplitude of the waves captured by the equipment (Hacki, 1996; Kania et al., 2004).

EGG analysis permits to verify the opening quocient (OQ) from the vocal folds. This measurement is related to the time proportion in which vocal folds remain opened during a vibration cycle (Verdolini et al., 1998). This measurement permits the clinician to verify the opening efficiency or inefficiency from the vocal folds (Weinrich et al., 2005).

Electroglottograph may be especially suitable for use in a clinician's office since it is immune to environmental acoustic noise, is portable, is relatively inexpensive, and software is commercially available for analysis of EGG signals (Cheyne, Nuss, Hillman, 1999).

Cheyne, Nuss, Hillman (1999) studied sustained vowel /a/ production in children aged from 3 to 16 years-old and concluded that the four measures on the EGG signal— jitter, OQ, closing quotient, and opening quotient—are independent of age and sex.

Aim: To analyze the opening quocient (OQ) from Brazilian Portuguese fricative voiced sounds produced by two children without communicative disorders.

Methods: one female subject aged 8:2 years-old and one male aged 7:5 years-old without communicative disorders complaint were evaluated. Both children are part of a bigger speech sample collected at the Laboratory of investigation in phonology from the Faculty of Medicine at the University of São Paulo. Children were submitted to a complete speech and language evaluation and results did not demonstrate any kind of disorders.

Eletroglottographic data (EGG) were collected by soliciting to the children a three-time repetition of six vehicle-sentences containing three target fricative voiced sounds from Brazilian Portuguese (Table-1).

Table 1: Vehicle-sentences used for repetition of the target words (in bold):

Target sound	Vehicle-sentence
/v/	Fale **vaca** bonita Say beautiful **cow**
/v/	Fale **vila** bonita Say beautiful **village**
/ʒ/	Fale **giz** bonito Say beautiful **chalk**
/ʒ/	Fale **jato** bonito Say beautiful jet
/z/	Fale **doze** bonito Say beautiful **twelve**
/z/	Fale **zelo** bonito Say beautiful zeal

Speech data was collected in a silent place using Computerized Speech Laboratory 4300B (CSL-4300B) and further analysis was made using CSL-4500. Data were analyzed by measuring target-sound duration (in seconds), opening quocient-OQ- mean (in percentage), minimum and maximum values from opening quocient of the target-sound (in percentage) and standard deviation-SD- (in percentage).

Results and Discussion:

Table 2 presents mean values obtained after three repetitions from the vehicle sentence, containing target words and sounds.

Table 2: Mean values for the eletroglottographic parameters.

	Subject 1-female					Subject 2-male				
Target words	Duration (sec)	OQ* (%)	Minimun (%)	Maximun (%)	SD*(%)	Duration (sec)	OQ* (%)	Minimun (%)	Maximun (%)	SD*(%)
vaca	0,10	51,85	32,92	64,63	9,37	0,11	41,14	39,25	44,07	1,39
vila	0,10	58,39	45,48	69,15	6,34	0,13	42,23	36,11	58,06	5,81
giz	0,08	53,81	33,51	66,51	9,97	0,12	44,46	40,55	48,12	2,05
jato	0,09	52,63	30,87	68,52	11,50	0,10	40,57	35,87	44,38	2,33
doze	0,09	58,97	56,02	63,14	2,13	0,10	43,74	42,04	45,28	1,06
zelo	0,11	51,10	43,96	58,43	4,04	0,11	41,96	32,54	47,69	3,91

OQ- Opening Quocient; SD- Standard Deviation

In general terms results demonstrated that duration of the three voiced fricative sounds from Brazilian Portuguese were practically the same for the target words in vehicle sentences

presented. OQ percentage for subject 1 was always higher than the same percentage for subject 2. The same results were observed for maximum OQ and SD mean values.

Another important result can be seen by comparing the OQ values of each target voiced fricative sound. Mean percentage of OQ values were similar to the three sounds independent of the subjects gender. This result suggests that voiced fricative sounds from Brazilian Portuguese tend to be produced using a similar percentage of vocal folds opening and they do not suffer interference from articulatory point.

Conclusion:

Eletroglottographic analysis from the OQ values of Brazilian Portuguese fricative voiced sounds demonstrated to be a reliable measure that can be used in order to contribute for a better characterization of Brazilian Portuguese sounds. Further studies using this parameter shall be made especially as an attempt to verify possible reasons of why devoicing phonological process is so occurring in Brazilian Portuguese.

Analysis of voiced and voiceless consonant production in children between 6 – 8 years old with and without speech disorders

Lucila Rey Rocha Schliemann.
schliemann@terra.com.br

Some children are sent to speech therapy because they substitute voiced sounds – /b/, /d/, /g/, /v/, /z/ e /Z/ – for voiceless sounds – /p/, /t/, /k/, /f/, /s/ e /S/ (e.g., “vaca” vs. “faca”). This phenomenon occurs in speech, writing or both. This study will address the problem in speech. It is well known that voicing is produced by differences in subglottal and supraglottal pressure.

A few studies have approached the problem in Brazilian Portuguese (Levy, 1993; Bonatto, 2007; Wertzner, Pagan, Gurgueira, 2009). Some have uncovered individual differences in the speech production of children with speech disorders. More specifically, they analyze the timing of the voicing contrast. The aim of this study is to examine whether the phenomenon is really a sound substitution or whether each child has different strategies to approximate the production expected in standard language.

Currently, speech clinicians work on the assumption of that voiced sounds are substituted for voiceless sounds. By means of the hypothesis that children with speech disorders try to approximate the normal production pattern, this study seeks a new way to evaluate and treat the problem in practice.

Six boys between 6 and 8 years old were selected. Three were diagnosed with speech disorders (they substitute voiceless for voiced consonants) and three presented normal speech. Initially speech screening tests were performed, which consisted of spontaneous speech, picture naming and Audiometry. The difference in speech production between the two groups yielded on the contrast of voicing. Hearing tests yielded normal results.

We selected 36 words for the *corpus*. Each word had voiced and voiceless consonants associated with the vowels / a /, / i / and / u /. The words were disyllables with the target consonant – voiced or voiceless – at the onset of the stressed syllable (e.g., “vaca” vs. “faca”).

First we played a game to find the pictures on a board. Such pictures referred to the words in the *corpus*. The children thus had a first contact with them. After the game, each word was inserted in the carrier phrase, “I found –WORD– pretty quickly”. The children uttered each phrase 5

times. The pictures were presented in random order. The recordings were made with a Plantronycs headset microphone in an acoustic booth with a Handy digital recorder.

The acoustic analysis was performed using Praat, version 5130. The words with voiced and voiceless consonants were segmented according to the followings variables: closure, VOT and burst (for plosives), friction (for fricatives), voicing (for voiced consonants), stressed vowel and second syllable.

We intend to measure the absolute and relative duration of each variable. Afterwards, we will analyze the vowel formants.

Statistical tests have not been performed yet. Up to now, we have observed variations in production within both groups. It seems that there are individual differences in speech production for each child. This is in line with previous work (Levy, 1993; Wertzner, Pagan, Gurgueira, 2009).

Articulatory Phonology (Browman and Goldstein, 1992, 1995; Goldstein, Byrd and Saltzman, 2006) has described such gradient processes. It has theoretical assumptions that may explain them. We hope to work out an account in which it throws light on the disordered children's strategies to get around their voicing difficulties.

References

BONATTO, M.T. Vozes Infantis: a Caracterização do Contraste de Vozeamento das Consoantes Plosivas no Português Brasileiro na Fala de Crianças de 3 a 12 anos. Dissertação de Doutorado. PUC/SP, 2007

BROWMAN, C.; GOLDSTEIN,L. Articulatory Phonology: an overview. *Phonetica*, 49, pp. 155-180, 1992.

GOLDSTEIN, L.; BYRD, D.; ELLIOT, S. "The role of vocal tract gestural action units in understanding the evolution of phonology" in *Action to Language via the Mirror Neuron System,* ed. Michael A. Arbib. Published by Cambridge University Press, 2006.

LEVY, I.P. A Nova face da nau dos incensatos. Tese de Doutorado. IEL/UNICAMP, 1993.

WERTZNER, H. F.; PAGAN, L. O. , GURGUEIRA, A. L. "Influência da Otite Média no Transtorno Fonológico: Análise Acústica da Duração das Fricativas do Português Brasileiro" in Revista CEFAC, vol. 13, 2009.

The influence of environmental noise on the performance of students in tests of tonal pattern of frequency and duration

Ludimila Nascimento, Stela Lemos and Rui Rothe-Neves.
ludimila.fono@gmail.com

Introduction:

Noise pollution is an environmental problem that has become known for its great influence on the physical and emotional health of individuals. In the school environment, in particular, it can cause damage to the process of teaching and learning by interfering in class activities. In speech therapy clinical practice, there are frequently cases of children with complaints of learning difficulties, which may be aggravated or caused by daily exposure to noise in the classroom. The evaluation of auditory processing in school is an important procedure for early detection and intervention in possible changes in auditory skills relevant to a good school performance. In this study, evaluation of auditory information processing was pursued in a natural environment of learning, that is, at school, thus aiming at assessing learners' capacities in the actual situation of demand, with difficulties inherent in their daily lives.

Purpose:

Characterize and discuss the performance of students in tests of tonal pattern of frequency and duration tonal pattern in a quiet and in noise environment.

Methodology:

This is a cross-sectional study in a convenience sample comprised by 70 students, aged from 10 to 15 years, 28 males and 42 females enrolled in one school in the public school system in Belo Horizonte, Brazil. Each student responded to a form with the purpose of mapping of noisy and visual analog scale to quantify the interference of noise, as well as their perception and causes. The noise level of the school was assessed by physical measurements with sound level meter. Class mates have been evaluated for speech, language, orofacial movements and simplified auditory processing. In the evaluation of speech we used phonemic context in which all phonemes were recorded as present, as well as substitutions, omissions or misstatements made during the utterance of phonetically balanced words. Language assessment was performed by applying logical and temporal sequence. To assess orofacial movements, we inspected palate, teeth and classified

occlusion; tension and mobility of phono-articulatory organs. There was also a simplified evaluation of auditory processing by means of sequential memory and verbal and nonverbal sound localization. Further research was made of cochlear-palpebral reflex (CPR) to 100 dB SPL to establish the presence or absence of hearing impairment. The absence of this reflex can be found in patients with

problems or conductive hearing loss. After performing these procedures, children were divided into two groups: no change in speech (G1) and patients with speech-language disorder (G2). G2 children failed in at least one of the evaluations (speech, language, sensorymotor and simplified evaluation of auditory processing) and G1 had adequate responses in all procedures. Children were then tested in standard tonal frequency and duration, with sequences of three or four elements, in two testing environments, in a quiet environment and in the presence of ambient noise.

Results:

Sound pressure level at the school ranged from 56.9 dB (A) in the computer room to 99 dB (A) in the schoolyard during recess. As for the form, children rated the noise coming from the schoolyard and the classroom itself as the most intense. According to children, the buzz of conversation and noise during class in this school was rated as high by 77.63% of respondents. There was a statistically significant association between environments for testing: three standard frequency sounds and duration pattern with three or four elements per sequence. The group without speech-language disorder (G1) performed better on tests of pattern tonal frequency and duration. In noise, performance of the two groups was worse.

That is, even subjects without speech problems had worst performance in noise, indicating the large interference of high noise levels in auditory processing of information. Difference between G1 and G2 was statistical significance in the tests: standard with four tonal frequency sounds in noisy environments and test pattern with three tonal frequency sounds, tonal pattern length of three or four elements, in a quiet environment. It was also found that there was a correlation between the perceived interference of noise and worse performance in tests.

Conclusion:

The sound pressure levels of the school are high and well beyond the recommended national standards. In noise, performance of tests in both groups was worst. That is, even subjects without speech problems has the worst performance in noise, indicating the large interference from high noise levels in the auditory processing of information.

$C_{nasal}V$ and $C_{stop}V$ biases in Brazilian Portuguese acquisition: an exploratory study

Magnun Madruga.
magnun.rochel@gmail.com

Speech acquisition data exploring CV co-occurrence patterns are generally investigated to account for the biological basis of the canonical syllables found in babbling and first words. The Frame-then-Content theory of the syllable (MacNeilage & Davis, 2000a,b) explored certain combinations of preferences, which are: 1) Labial consonants with central vowels; 2) Coronals with front vowels; 3) Dorsals with posterior vowels. The frequency counts come from oral language transcripts and are

used to corroborate the theory's main assumption, namely: biomechanical movements originated from mandible oscilations provide the frame for the canonical syllable, which will be followed by the content of a richer phonetic-phonological repertoire both in phylo and in ontogenesis. Based on this study, our investigation aims to check whether the CV biases reported by MacNeilage & Davis are found in CV combinations involving stops and nasals.

The data come from longitudinal samples of the Brazilian Portuguese acquisition - BP - of seven (7) babies whose ages range from 1:5 to 2:5. The data come from LIDES (Linguagem Infantil em Desenvolvimento) databases and are characterized by being collected in natural interactions between the babies and their caregiver(s).

After being transcribed orthographically, the data went through an orthographic-phonic converter, ORTOFON (ALBANO & MOREIRA, 1996), and, after converted, were quantified using the software WordSmith Tools, version 5.0 (SCOTT, 2008).

We investigated CV biases in syllables composed of stop and nasal consonants in three different word positions: a) initial unstressed b) medial stressed; and c) final post-stressed. We performed type and token frequency counts and ran a statistical analysis using the Chi-Square Test to verify the existence of a relationship between the point of consonant and vowel articulation. To check the strength of this association, we used Crammer's V, which measures the significance of the association. At last, we used the O/E Ratio, aiming to offer a qualitative view of the preferred CV biases found in the infants' data according to the position of the syllables in the words produced.

The findings from this study partially confirm the CV combinations MacNeilage and Davis

found, but only in some positions. For example, it is observed that in Portuguese, in the type frequency counts, the vowels have a strong association with labial consonants. This fact shows the evidence that vocabulary type used by children, who have highly frequent labial nasals and stop consonants reduplicated in their composition, have a strong effect on the emergence of phonic content. In (1), we analyze the O/E Ratios found in the three positions investigated.

Table (1)

O/E Ratio	Initial		Stress		Post-stressed	
	Type	Token	Type	Token	Type	Token
C labial - V central	0,85	0,88	0,68	1,64	0,68	0,75
C labial - V front	1,42	1,11	1,21	0,74	1,21	1,82
C labial - V back	0,95	1,04	0,59	1,38	0,59	0,72
C coronal - V central	0,89	1,21	1,49	0,28	1,49	0,88
C coronal - V front	1,16	0,87	0,77	1,27	0,77	0,91
C coronal - V back	1,10	0,78	1,22	0,63	1,22	1,43
C dorsal - V central	1,28	0,8	1,02	2,01	1,02	1,47
C dorsal - V front	0,43	1,03	0,90	0,62	0,90	0,31
C dorsal - V back	0,84	1,30	1,30	1,47	1,33	0,45

In (1) we can perceive the differing biases posed by type and token frequency counts. This suggests that the infant lexicon preserves certain patterns in types, but rejects them in their frequency of tokens or vice versa. This evidence can be found easily in the combination labial C-central V, which shows a strong bias only in stressed position. On the other hand, the combination labial C labial – front V shows opposite biases. These significant differences found in CV co-ocurrence patterns can be more adequately explained in the light of Dynamic Theories, which predict that the materiality of language, grammar, has a great influence on the distribution of these biases. Gestural Phonology allows us to refute MacNeilage and Davis's (2000a,b) argument that ontogenetic development would be recapitulating phylogenesis. Instead, it highlights the fact that the infant's development is subordinated to more general principles of dynamics. Gradually, with the development of vocabulary, children acquire biases that are imbricated in the linguistic system (ALBANO, 2009).

Finally, we conclude that the CV biases found in this study differ from those exhibited by MacNeilage and Davis. Therefore, it is necessary to explain them with a different theory. We argue in favor of dynamic theories, since they are able to explain the CV co-occurrence biases by principles such as the dynamics of the task and coupled oscilators, simply and economically. Moreover, it is pertinent to explain that the phenomenon is also linked to the accentual position,

since the biases shift according to the position of the CV syllable in the word.

References

Albano, E. C.; Moreira, A. A. *Archisegment-based letter-to-phone conversion for concatenative synthesis in Portuguese.* Proceedings ICSLP'96, 3, p. 1708-1711, 1996.

Albano, E. C. *A corpus-based reappraisal of the role of biomechanics in lexical phonotactics*. Oral Presentation in the V Corpus Linguistics Conference, 2009.

MacNeilage, P.F.; Davis, B.L. *On the origin of internal structure of word forms*. Science, 288, p. 527-531, 2000a.

MacNeilage, P. F.; Davis, B. L.; Kinney, A; Matyear, C. L. *The Motor Core of Speech: A Comparison of Serial Organization Patterns in Infants and Languages.* Child Development, vol. 71, nº. 1: p. 153-163, 2000b.

Scott, M. *WordSmith Tools version 5,* Liverpool: Lexical Analysis Software, 2008.

Production of [m] and [n] in Codas by Brazilian Students: an Acoustic Analysis

Márcia R. Becker.
marcia.r.becker2009@gmail.com

In the process of acquisition of a second language (L2), difficulties inevitably arise in the production of certain phones/sounds in its inventory. The production of the phonemes [m] and [n] in coda position can be considered as cases in which there are difficulties, if we take Portuguese as L1 and English as L2. The point is that Brazilian learners generally transfer to English the knowledge they have of the Portuguese phonic system and its orthographic conventions, tending to produce nasalized vowels, without making the distinction between [m] and [n] at the end of syllables.

The objective of the research was to identify through acoustic analysis what was produced by Brazilian pre-intermediate learners of English in codas of monosyllables which have the sequence vowel + nasal consonant ([m] and [n]), using minimal pairs whose contrast lies precisely in the nasal sound, with two distinct environments preceding the vowel ([æ] and [I]), and with silence after the nasal to avoid problems of co-articulation, very common with nasals. Acoustic analysis was preferred over auditory analysis because the contrast between final [m] and [n] is difficult to hear even for native speakers. The analyses were mainly comparative, with the use of the productions of two native speakers, who compose the control group. The production of ten pairs of monosyllabic words (minimal pairs) was recorded. The informants were ten adolescents, students of English at pre-intermediate level, five boys and five girls. Three repetitions of each word of the *corpus* by each of the informants were recorded, making 600 words for the students and 120 words for the control group. Data were analyzed using PRAAT, version 4.4.30.Because of the lack of specific references about acoustic data of nasals in codas, the analysis of the data obtained with the control group was much more detailed, so that we could, based on the results of such analysis, define one or more parameters for the analysis of the production of the Brazilian learners.

After analyzing the relative duration of the nasal consonants (obtained by dividing the average value of the duration of the nasal murmur by the total duration of the target word), the three first formants of the vowel preceding the nasal, and the three first formants of the nasal (specifically the nasal murmur) of the data from the control group, the values of the second formant (F2) of the nasal proved to be the most consistent ones. For all the words in the *corpus*, in each of the three

repetitions, for both informants, the figures related to F2 were consistently higher for [n] than for [m], for the same pair. This was therefore the parameter used to distinguish between [m] and [n]. Ladefoged had already observed the fact that, when the lips close to produce [m], the second formant of this consonant lowers its frequency (Ladefoged, 2001, p. 54), but he did not make a comparison with the production of [n].By comparing the learners' productions with those of the native speakers, it was shown that 40% of the target words produced by the learners were acoustically similar to those produced by the native speakers, 27% of the words showed an F2 with opposite results to those of a native speaker, while a third of the productions were acoustically indistinguishable (F2 of [m] and F2 of [n] showing very close frequencies, though F2 [n] was higher than F2 [m]).In relation to the preceding vowels, 44% of the words with [I] produced by the learners were similar to the words produced by the native speakers, whereas only 36% of the words with [æ] proved to be similar.

Very few studies mention nasal consonants in the interlanguage of Brazilian learners of English. The only acoustic study found (Monahan, 2001, p. 23) points out that "BP regressively assimilated nasalization onto a vowel from a nasal consonant in coda position and deleted the nasal consonant". Kluge (2004), in experiments of production and perception of nasal consonants (by ear), came to the same conclusion. So, One of the main purposes of this study was to define a parameter which could be used acoustically to identify whether a production was a segment [m] or [n] in a coda, since the analysis by ear of bilabials and alveolars in codas is very controversial. This aim was achieved, bearing in mind the limited conditions of this research: F2 of [n] proved to be always higher than the F2 of [m].

References:

Kluge, D. C. (2004) *Perception and Production of English Syllable-Final Nasals by Brazilian Learners.* Florianópolis, 77f. Dissertation (Masters in English and Literature), Universidade Federal de Santa Catarina.

Ladefoged, P. (2001) *Vowels and Consonants:* An Introduction to the Sounds of Languages. Malden, MA: Blackwell.

Monahan, P. J. (2001) *Evidence of transference and emergence in the interlanguage.* Retrieved 05 August 2007, from http://roa.rutgers.edu/files/444-0701/444-0701-MONAHAN-0-0.PDF.

Nasality as an acoustic parameter for speaker identification

Márcio Pinto
oppliger@terra.com.br

Forensic identification techniques such comparing DNA or fingerprints are founded on the uniqueness features that these elements present or, in other words, on the idea that the best prove in identification process is the evidence that shows low intra-subject variation (we can assume that there is almost no variation of two fingerprints or DNA samples of the same person) and relatively high inter-subject variation (we can also assume that there is no different persons sharing the same DNA – with the exception of identical twins – or fingerprints).

The acoustic parameters commonly used in forensic speaker identification do not demonstrate this relation between inter and intra-speaker variation (or "within-speaker" and "between-speaker" variation [1]), since an individual's vocal tract does not present the same configuration on each utterance of the same word, even when we try to maintain the same context conditions around the word.

The vocal apparatus is probably the most important source of phonetic differences between speakers, since the constitution of the human vocal tract varies according to the anatomy and physiology of each individual. But, besides this close relation with anatomy, the vocal tract will not be configured just by biological (and fixed) factors. Mobile articulators change their relative positions and shapes to produce speech sounds. So this dynamic modelling function of the vocal apparatus establishes constraints which determine a range within an acoustic parameter takes place [2]. It is in this scenario that vowels' formants values are the most used acoustic parameters in the speaker identification process.

We can expect that acoustic parameters related to the nasal tract would offer better profile in comparison task, since the configuration of nasal cavity, as a fixed structure, does not suffer alterations in its shaping, and, consequently, acoustic measures of that cavity would present less within-speaker variation than the vocal tract ones.

It's not the tradition within the Forensic community the use of nasal acoustic measures as comparison parameters, even though these measures had been indicated as better parameters by the literature [3]. Studies related to the nasal cavity's acoustic behaviour will offer a better understanding on the nature of such resonator and its implications in forensic field.

The participation of nasal cavity as a resonator in the speech process is independent of the participation of the oral cavity, though articulations that involves nasal features could show in some extent the overlapping, in time, of nasal and oral parts.

Traditional segmental models could not satisfactorily explain overlaps of these phenomena, since the linear notion of the processes limits the theory scope. Task dynamics, otherwise, seems explain it in a better way, offering an interesting view of co-articulatory influences between sounds and gestures.

Task dynamics can maybe offer a better understanding on the articulatory facts involved in nasality and its importance as an acoustic parameter for the speaker identification process.

REFRENCES:

[1] Rose, P. (2002) *Forensic Speaker Identification,* London: Taylor & Francis.

[2] Nolan, F. (1983) *The Phonetic Bases of Speaker Recogni tion*, Cambridge:
Cambridge University P r e s s .

[3] WOLF, J . J . (1972) Efficient acoustic parameters for speaker recognition. *Journal of the Acoustical Society of America*, v.51 n.6, p. 2044 - 2056.

The Emergence of Some Phonological Patterns in the Verbal Morphology in Brazilian Portuguese

Marco Fonseca and Mário Silva.
marcosilvafonseca@gmail.com

This research aims to analyze the emergence of some phonological patterns in the verbal morphology of Brazilian Portuguese (BP) which are related to current language variation and change (Naro, Scherre 2003). The case studies to be considered involve segmental loss and the focus of the analysis is mainly on the mechanisms which lead the emergence of new patterns. For example, a verbal form such as *falaram* [faˈlaɾãu] '(they) spoke' occurs as [faˈlaɾu], where reduction is observed in the loss of postonic nasality.

The analysis to be presented is based on Usage-Based Phonology (Bybee 1985, 2001). Exemplar Models will provide the representational characterization of the findings to be presented (Pierrehumbert, 2001). The main point to be addressed is the role played by type and token frequency in relation to language variation and change in the verbal morphology in the cases to be analyzed. It will also be shown that a network model is fully adequate to handle elegantly the case studies we have considered.

Usage-based models consider frequency to be a key factor to determine productivity. There is type and token frequency. Token frequency refers to the occurrence of a certain unit in a given corpus and determines the lexical strength of a given pattern. Type frequency, on the other hand, refers to the dictionary count of a certain pattern (Bybee 1985, 2001). The analysis to be presented considered cases of segmental deletion in BP's verbal inflections, as mentioned above. The data to be analyzed were qualitatively evaluated according to two sources. Data were formerly analyzed according to Brazilian Portuguese traditional grammars with the aim to identify regular or prescriptive phonological patterns. At this stage potential cases of sound variation and change were identified. This step was followed by an evaluation of data from Projeto Aspa (www.projetoaspa.org), which provides phonological type and token frequency values from a BP corpus. Then, potential cases of observed phonological changes were evaluated according to frequency effects. The main hypothesis of this study is that segmental loss is closely related to the prosodic organization of the language at all levels. Thus segmental loss affects mental representations. In other words, it will be argued that prosodic reorganization follows from

frequency effects observed in the language in question. Therefore, segmental loss might be understood as an accommodation of more frequent or regular phonological types. However, frequency effects are subjected to structural factors and also to general usage of the language. This predicts that there might be sets of structurally adequate forms that do not undergo the expected sound change. It also predicts that specific language usage may preserve forms that should have undergone the expected sound change.

The results suggest that high token frequency contributes to maintain the irregular forms. Frequently used irregular verbs, such as *ser* 'to be' and *ir* 'to go' comonly involve root suppletion and other irregular patterns, but their stability is secured by great lexical strength due to their high token frequency. For example, the verbal form *era* 'was' has a high token frequency value of 201,845 in Projeto Aspa database. We demonstrate that high type frequency is responsible for regular verbs to act like a magnet, atracting to their paradigm the low token frequency irregular verbs (Phillips 2001). Our analyzis indicates that irregular forms which presented low token frequency tended to regularize due to phonological leveling or generalisation. For example, verbs such as *romper* 'to rupture' tend to lose their irregularity and to fit into high type frequency regular patterns observed for regular morphology as a consequence of type frequency effect. Thus, instead of having the irregular form *roto* (token frequency value of 130) it starts showing the regularized form *rompido* (token frequency value of 847).

The results indicate that verbal tenses of higher token frequency displayed a higher rate of deletion. Thus, higly used items as *ficar[ãu]* → *ficar*[u] '(they) stayed' or *gosta[mus]* → *gosta[mu]* '(we) like or liked' presented a higher rate of deletion than low frequency verbs such as *estocam* '(they) store' or *estocaram* '(they) stored' (token frequency values of 30 and 14 respectively) which indicates an important role played by token frequency on the implementation of this specific change. Type frequency also plays an important role in sound change. Thus, although s-deletion might be expected in forms such as *chega[mus]* → *chega[mu]* '(we) arrive or arrived' it is hardly observed in other verbal conjugations such as *escrevêra[mus]* → **escrevêra[mu]* '(we) had written'. This indicates that it is not simply s-deletion that takes place, but in fact s-deletion takes place under specific lexical conditions. Another case to be considered is the reduction of diphthongs in the 1st conjugation *cham[ou]* → *cham[o]* '(he/she) called', which does not occur in verbs of the 2nd and 3rd conjugation (*venc[eu]* → **venc[e]* '(he/she) won') because verbs of the 1st conjugation exhibit higher token frequency. However, other factors such as phonetic similarity and stress should be considered. The three case studies shown above indicate that the similar segmental sequences behave differently in relation to segmental reduction or deletion.

In conclusion, the results indicate that the investigated cases of phonological reduction

apply to specific morphological classes, i.e., although they have apparent identical phonological properties some cases undergo reduction and others do not. This suggests that morphology is organized in a network fashion (Bybee 2001). We argue that exemplar models are the most suitable framework to explain these findings. A final remark should be made in relation to the fact that segmental deletion observed in high frequency verbal tenses lead to the emergence of -CV final syllables in verbal morphology of BP. However, other phonological phenomena indicate some segmental changes towards more complex syllables rather that CV ones. For example, *pa[htisip]ar* → *pa[htsp]ar* 'to participate' presents a complex coda or *[distr]ibuído* → *[dstr]ibuído* 'distributed' shows a complex onset. Future research should attempt to understand the balance between the forces that trigger reduction towards a CV syllable type whereas other types of change trigger the emergence of complex syllables. Another issue to be investigated in future research concerns the prediction of Usage-based Phonology that regular sound change is phonetically gradual.

References

- Bybee, J. (1985) Morphology – a study of the relation between meaning and form. John Benjamins, Amsterdam/Philadelphia.
- Bybee, J. (2001) Phonology and Language Use. Cambridge Univ. Press, Cambridge UK.
- Naro, A. J. ; Scherre, M. M. P. (2003) Estabilidade e mudança lingüística em tempo real: a concordância de número. In: M. C. Paiva; M. E. Duarte. (Org.). Mudança lingüística em tempo real. Rio de Janeiro: Contra Capa. 47-62.
- Phillips, B.S. (2001) Word-frequency and the actuation of sound change. Language 60. 320-342.
- Pierrehumbert, J. (2001) Exemplar dynamics: Word frequency, lenition, and contrast. In J. Bybee and P. Hopper (eds.) Frequency effects and the emergence of lexical structure. John Benjamins, Amsterdam. 137-157.

Frequencies Scales: an intonation and perception's speech approach

Marcus Vinicius Moreira Martins.
marcusvmmartins@gmail.com

Works involving pitch perception or speech intonation usually uses the international standard scale Hertz (Hz) to plot the data obtained in the measurement of F0 curves; however, we can trace in literature the use of some other frequency scales, like the semitones (st), mels, barks and even ERB-rate. Thus, the present study, which is part of a broader project, aims to examine what, would be the most appropriate frequencies scale for data analysis in studies involving intonational aspects of Brazilian Portuguese (BP), regarding mainly speakers perception and intuition at the time of the production.

First is necessary to remark an important point: almost all of these scales come from mathematical functions pre-established by psychological experiments (see Hermes et al., 1991), thus they establishes relations with the Hertz scale. In this sense, as theoretical and methodological approach, we chose compare the relation between Hertz (Hz) and semitones (st) in this first study. We take this position due to two important aspects: the first one is because we understand, based on Roederer (2003), the chromatic scale as the most widespread and "natural" frequency scale in the Western culture with regard to the values of heights. The second factor is that the semitones scale is a Hertz's logarithmic transformation, it is: a geometric progression of ratio q = 1.0594 (st1 = f1 x 1.0594 – almost 6%), thus, we operate with frequencies ranges and the hypothesis that there is a center frequency for where the neighboring frequencies, in a given range, are attracted. In this sense, we analyses the relative distances between frequencies instead of the absolute frequencies themselves, justifying the use of the scale of st ('t Hart et. Al, 1990).

Experiment and analysis

We based our experiment in a study performed by Francis Nolan (2003). We proceeded in the same way, but in our test we prefer work with an utterance taken from a movie scene, which have two turns, being, respectively, a man and a woman. Our choice is justified by the utterance's prosodic features as can be seen below in (1):

(1) **M:**[Eu tava dormindo…][de repente:: acordei…][e dei de cara com o Domenico.]

L H% H L+H% H L%

F:[O Domenico entra aqui::?][como assim?][1]

L H% H%

The recordings were madding in a noise-controlled environment, using a unidirectionaldigital recording microphone. The test was performed with 13 BP native speakers (being 7 women and 6 men) and consisted primarily in an exposure to the target-utterance (1). In next step, they were required to reproduce what they had heard: imitating just the intonation a despite of the voice qualities and segments realization, in the sense that they have not to imitate the dialectal variation. During the test, the same subject reproduced the both voices, a despite of the sex, our intention was to assess if there is any relation in the interpretation of voices with lower F0 or higher F0.

The recordings obtained were analyzed in ExProsodia routine which got all the prosodic features of the recordings, including the syllabic values of F0 (the average of valid values of F0 in the syllable) in st (MIDI) and HZ, what were the basis for our analysis. Then we calculate the correlation (r2), using Microsoft Excel, between the values of syllabic F0, it is we compare the thirteen recordings one-by-one with the utterance-target in both scales (st) and (Hz). Thus, if the value of the correlation's difference between scales was relatively large, we could say which is the best frequency scale. In this way, we obtained four r² averages, two Hz-st for female voice and two Hz-st for male voice as can be seen in Table 1:

	Female Voice	**Male Voice**
Hertz (HZ)	$r^2 = 0{,}639012$	$r^2 = -0{,}06466$
Semitones (st)	$r^2 = 0{,}62283$	$r^2 = -0{,}06455$
Variation (r^2Hz-r^2st)	0,016182	-0,00011

Table 1.R ² values obtained in the Hz-st male and female correlation.

We may note in the first set of data (female voice) that the variation of correlations is around 0.01 and cannot be regarded as significant. In the second data set (male), we note that the variation is around 0.0001, what also represents a not meaningful variation. Therefore, we cannot say which of these scales are more efficient to represent the processes of perception and intuition of

the speakers, although it is plausible to say, according to the data, both can be used for the same application without systematic losses in the data quantifying. It is also interesting to note that the speakers had a greater tendency ($r^2 \approx 0.6$) to imitate the target speakers imitate voices with a lower F0 whatever the frequencies scale. In the next steps, we intend evaluate others prosody's structural elements, as well as other scales, like barks and mels, we also intend evaluate the prosodic segmental features and others statistics analysis.

Bibliography

Roederer, J. (2003)Introdução a Física e a Psicofísica da Música. São Paulo: Edusp.

Nolan, F. (2003) Intonational equivalence: an experimental evaluation of pitch scales. Proceedings of the 15th ICPhS, Barcelona, pp. 771-774

Hermes, D. & van Gestel, J. (1991).The frequency scale of speech intonation. Journal of the Acoustical Society of America, 1991, 90, 97-102.

Hart, J. 't.; Collier, R. & Cohen, A (1990). A perceptual study of intonation: an experimentalphonetic approach to speech melody Cambridge: Cambridge University Press.

Notes:

[1]M:[I was sleeping…][suddenly::: I woke up…][and…I bumped with Domenico]
F: [Domenico come in here?] [How so?]

An experimental study on the perception of Brazilian Portuguese lexical stress

Maria Cantoni.
mmcantoni@gmail.com

It remains to be fully accounted for how rhythmic patterns are processed and stored in representation and their relation with linguistic structure. A point of great interest to this issue is the active role of acoustic parameters and gestural configurations involved in both production and perception of rhythmic patterns. This paper reports on preliminary results from a study that investigated acoustic and articulatory parameters employed in the perception of lexical stress in Brazilian Portuguese (BP). The analysis developed is based on experimental data and relies on complementary principles borrowed from usage-based [1] [2] and gesture-based models [3] [4], cognitive linguistics [5] and exemplar theory [6] [7].

In BP, lexical stress is presented as acoustic prominence of one of the last three syllables in the word domain. Traditional analyses assume that stress is contrastive, on the base of a few minimal pairs such as *cara* "face" vs. *cará* "yam". In this study, we intend to address the representational issue of lexical stress on the basis of its functional behavior, with support of experimental data.

On the past years, the study of rhythm from a phonological and prosodic point of view has been mainly devoted to the description of intonational systems from languages around the world, with the special aim of identifying language-specific and universal patterns and offering a rule-based model of rhythmic phenomena, usually separating production and perception as independent fields. Among the studies concerned with the production of rhythm, many of them were based either on impressionistic data or in acoustic analysis alone; lesser attention was payed to the intricate issue of how is rhythm physiologically implemented. In the side of perception, experimental approaches were explored, but mainstream researches were based on psychological grounds, usually adopting tones and non-linguistic stimuli.

With the emergence of usage-based and cognitive models in phonology, the use of corpora and experimental data has been increasingly pursued, as well as the need to integrate production and perception in a comprehensive model of language. A further development specially related to cognitive models is the view of language not as an innate and independent module, but rather as an

integral part of high cognitive processes, therefore fully embedded in the human mind, body and environment. As a consequence of this new paradigm, interdisciplinary work has been strongly emphasized, with the additional advantage of bringing together advances from varied areas also compromised with the study of human language, such as engineering, physics, neurosciences, psychology, artificial intelligence.

Exemplar theory claims that linguistic representations are multidimensional and web-like connected, challenging the uni-dimensional basis of representations proposed by traditional models. Specifically with regard to phonological theory, a remarkable depart from traditional theories has been the idea that the building blocks of speech are articulatory routines, or gestures, that result from the spatiotemporal coordination of articulators, as fully developed in gesture-based models. It follows from this assumption that nonlinear sequences of segmentsized and syllable-sized gestures are perceived and produced as whole and unique articulatory routines. Based on the aforementioned theories, we assume that segments and syllables are not independent units of speech; rather, as they are coordinated in time with suprasegmental units such as stress and intonation contours, they emerge as general patterns across conceptually grounded larger units, most likely words, but also chunks and collocations.

The experiment was divided in two parts, one based on non-linguistic stimuli, and the other based on linguistic stimuli. The first (non-linguistic) part was inspired in Bell [8] and aimed at extending his study to Brazilian Portuguese. As in the original experiment, four kinds of three-tone sequences were used: a control sequence, in which the tones had all the same characteristics; and three test sequences, in which one of the three tones of the sequence was made stronger in either intensity, frequency or duration. The subjects were told to choose the best match to the auditory pattern they heard, regarding the order of the strong (s) tone in relation to the weak (w) tones: s-w-w; w-s-w; w-w-s. Bell, whose experiment tested six languages with different rhythmic systems, concluded that the perception of non-linguistic rhythmic patterns was not significantly influenced by specificities of the speaker's native language. Contrarily, the results of our study on BP were in accordance with Jakobson, Fant and Halle [9], pointing to tendencies that could be related to influences of BP sound system, such as a prevalence of the w-w-s and w-s-w patterns in frequency and intensity sentences.

The second (linguistic) part was similar to the first one, but used linguistic stimuli instead. Tones were replaced with pairs of words whose continuous repetition could lead to ambiguous interpretation (i.e. with first or last syllable displaced, e.g. *marrom* [ma'hõ] “brown” vs. *Roma* ['hõma] “Rome”). The words were chosen with regard of their token frequency, in such a way that some pairs had similar token frequency values, while others had dissimilar values. As expected,

results pointed to token frequency effects over the perception of ambiguous syllable sequences. Token frequency would influence the choice within the pairs with dissimilar token frequency values. On the other hand, in pairs with similar values, token frequency effects were obviously null and type frequency influenced to some extent the choice between pairs. This test also served the purpose of discussing whether linguistic and non-linguistic rhythmic patterns are similarly perceived. Based on these and other experiments, we concluded that studies using non-linguistic stimuli are to be carefully interpreted, always bearing in mind their limitations, such as the lack of a usage variables and conceptual grounding.

The experiments presented above are part of a broader project devoted to investigate BP rhythmic patterns from a cognitive perspective, considering both perception and production. To date, we are designing new perception tests to address other issues regarding BP rhythm. Their results will be compared with results presented here as well as with results from a previous production experiment. We intend to examine whether perception and production involve the same acoustic parameters, and whether such parameters have the same weight. With the analysis to be developed, we intend to investigate to what extent usage- and gesture-based models, cognitive phonology and exemplar theory can be brought together in a complementary way to achieve a comprehensive and robust model in phonology and prosody.

References

[1] Barlow, M.; Kemmer, S. (ed.). 2000. *Usage-based models of language*. Stanford: CSLI.

[2] Bybee, J. 2001. *Phonology and Language Use*. Cambridge: Cambridge University Press.

[3] Browman, C.; Goldstein, L. 1986. Towards an articulatory phonology. *Phonology Yearbook*, 3, 219-252.

[4] Byrd, D.; Saltzman, E. 2003. The elastic phrase: dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, 31, 149-180.

[5] Langacker, R. 1987. *Foundations of cognitive grammar.* Stanford: Stanford University. [6] Nosofsky, R. 1988. Exemplar-based accounts of relations between classification, recognition, and typicality. *J. of Experimental Psychology: Learning, Memory, & Cognition*, 14, 700-708.

[7] Pierrehumbert, J. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In Bybee, J.; Hopper, P. (ed.). *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins.

[8] Bell, A. 1977. Accent placement and perception of prominence rhythmic structures. In Hyman, L. (ed.). *Studies in stress and accent*. Los Angeles: University of Southern California.

[9] Jakobson, R.; Fant, G.; Halle, M. 1952. *Preliminaries to speech analysis*. MIT.

Secondary Stress in Catalan

Marianna Nadeu.
mnadeur2@illinois.edu

Descriptions of stress in Catalan usually mention the existence of secondary or rhythmic stress (SS), apart from contrastive primary stress. Yet, these descriptions ([1]-[5]) are varied and there is no agreement on whether the phenomenon is obligatory or optional, whether it is purely rhythmic or morphologically-conditioned, and what its pattern is (binary or ternary feet). A few studies ([6]-[8]) have tried to find perceptual or acoustic evidence for its existence in morphologically simple and complex words with no success. However, a common assumption in these three studies is that SS is predictable in Catalan and that the initial syllable bears it. Their inability to find evidence for SS may be due to the fact that the materials they were analyzing did not contain any instances of it.

Hualde ([9]) provided examples of two types of optional SS in Spanish, which he calls 'rhythmic' and 'emphatic', and claimed that SS is a very recurrent feature in the language of the media, lectures, recitation and public speech in general. In addition, the existence of such varied proposals (for both languages) for what the typical patterns might be suggests that there is no clear intuition as to whether SS occurs invariably and whether it occurs in a specific position. In a production experiment, Hualde ([10]-[11]) succeeded in eliciting real examples of SS in Spanish. He found that syllables with SS were characterized by higher f0 and higher intensity, but shorter duration than the syllable with primary stress. Studies on SS in Brazilian Portuguese report similar findings. Unlike in Catalan and Spanish, in Brazilian Portuguese SS might be considered an inherent property of the prosodic word, due to its high frequency of occurrence. Abaurre and Fernandes ([12]) asked 17 speakers to read a magazine article. 97% of the prosodic words with one or two pretonic syllables were perceived (by three judges) to have SS. Binary feet were the most common pattern in their data, although ternary feet were also found in emphatic readings. The presence of a SS was also marked by a high tone. Previous studies had found that duration was not a cue of SS (cited in [12]).

Catalan remains unexplored in this respect. An important phonological feature of Catalan that is not shared by Spanish is the systematic reduction of the vowel inventory from 7 to 3 vowels in unstressed syllables, including a vowel schwa that is only found in that context. My goal is to

identify the acoustic correlates of SS as well as its patterns. In addition, I want to test the effects of SS on the quality of lexically unstressed vowels. My hypothesis is that whereas the presence of secondary, phrase-level, stress may result in less "phonetic" reduction it will not affect "phonological" reduction. In particular, the vowel schwa that appears in lexically unstressed syllables will not merge with any of the vowels in the stressedsyllable inventory. I propose this as a test of the "phonetic" (gradual) vs. "phonological" (categorical) character of both vowel reduction and SS. Comparison among the three Ibero-Romance languages, Catalan, Spanish and Portuguese, may shed much light on this issue.

I collected data from a Catalan radio news program (10 speakers: 5 males, 5 females, ages: 30-50). I will examine duration, f0, intensity and formant structure of the vowels /i/, /u/ and schwa in four types of syllables: 1) unstressed, 2) primarily-stressed, 3) secondarilystressed, and 4) in primarily-stressed syllables in words which also have a SS. Schwa will only be analyzed in contexts 1 and 3, the only ones in which it can occur. I will also determine the patterns of SS. More specifically, I want to test the traditional claims that characterize SS as creating an alternating binary or ternary sequence.

In a previous pilot study, I analyzed 103 instances of SS from data of the type described above. The most common pattern found was a binary one (87% of the cases), although SS could appear up to 4 syllables away from the primary stress. In fact, two important tendencies were noted: binary pattern and word-initial SS, which could point to two different, but related, phenomena ([9]). Regarding the acoustic correlates, primarily-stressed vowels were found to be significantly longer than unstressed and secondarily-stressed ones.

On the other hand, vowels with SS were found to have higher mean intensity than those with primary or no stress. Furthermore, a within-word analysis of f0 showed that vowels with SS had significantly higher mean f0 than those with primary stress in the same word. The results of this pilot study agree with findings for Spanish ([10-11]) and Brazilian Portuguese ([12]) in that the presence of a secondary prominence tends to result in an alternating, binary sequence.

Additionally, intensity and f0 were found to be correlates of SS in Brazilian Portuguese and Spanish, and they are in Catalan too.

The pilot study showed that it is possible to find acoustic evidence for the existence of SS in a specific type of data: language used in the media, which tends to be very emphatic and has a very idiosyncratic prosody. I hypothesize that this phenomenon is employed as a strategy to emphasize certain parts of the discourse and that it may involve a more careful articulation that can be related to Gussenhoven's *effort code* ([13]). Finally, I will speculate on the modeling of lexical and phrasal stress effects from the perspective of gestural organization of words and utterances.

References

[1] Coromines, Joan. 1974. Lleures i converses d'un filòleg, 2nd ed. Barcelona: El pi de les tres branques.

[2] Oliva, Salvador. 1977. Dos aspectes del ritme en el català modern. Els Marges 9. 89-96.

[3] Recasens, Daniel. 1993. Fonètica i fonologia. Barcelona: Enciclopèdia Catalana.

[4] Bonet, Eulàlia & Lloret, Maria-Rosa. 1998. Fonologia catalana. Barcelona: Ariel.

[5] Prieto, Pilar. 2002. Entonació. Models, teoria, mètodes. Barcelona: Editorial Ariel.

[6] Mascaró, Joan.1983. Apèndix I: Nivell perceptual de l'accent no principal en els compostos. In La fonologia catalana i el cicle fonològic. Bellaterra: Publicacions de la Universitat Autònoma. 197-202.

[7] Prieto, Pilar. 2003. Correlats acústics de l'accent secundari en català. Estudios de fonética experimental 12. 106-142.

[8] Recasens, Daniel. 1991. Timing in Catalan. Actes du XIlème Congrés International des Sciences Phonétiques, vol. 4. Aix-en-Provence. 230-233.

[9] Hualde, José Ignacio. 2007. Stress removal and stress addition in Spanish. Journal of Portuguese Linguistics 5/6. 59-89.

[10] Hualde, José Ignacio. 2009. Secondary stress Spanish. Paper presented at the 4th Phonetics and Phonology in Iberia Conference, Las Palmas, Canary Islands, Spain, June 17-19.

[11] Hualde, José Ignacio. In press. Secondary stress and stress clash in Spanish. To appear in Proceedings of Laboratory Approaches to Spanish Phonology 4, ed. by Marta Ortega-Llebaria. Cascadilla Proceedings Project.

[12] Abaurre, Maria Bernadete M. & Fernandes, Flaviane R. 2008. Secondary stress, vowel reduction and rhythmic implementation in Brazilian Portuguese. In L. Bisol, C.R.M. Brescancini (eds.), Contemporary Phonology in Brazil. Cambridge: Cambridge Scholars Publishing.

[13] Gussenhoven, Carlos. 2002. Intonation and interpretation: Phonetics and Phonology. In B. Bel & I. Marlien (eds.): Speech Prosody 2002: Proceedings of the First International Conference on Speech Prosody. Aix-en-Provence: ProSig and Université de Provence Laboratoire Parole et Language. 47-57.

The dynamics of interlinguistic transfer of VOT patterns in multilingual children

Marta Helena Bandeira.
martatbandeira@gmail.com

The small town of Arroio do Padre, located in south of Brazil, was colonized by Pomeranian and German immigrants and is characterized as a geographic region of relatively limited access, which contributes for the daily use of Pomeranian and/or German as the first language by the inhabitants of this city. But what phonetic-phonologic characteristics permeate the transfer across the languages spoken by a multilingual individual? Departing from this initial question, the aim of this study is twofold: 1) to analyze the dynamics of transfer in the VOT patterns of the initial obstruents /p/, /t/ and /k/ in the three languages spoken by multilingual children (within group analysis); and 2) to compare the VOT patterns of voiceless stops produced in English as an L2 by monolingual speakers of Brazilian Portuguese (BP) in the initial stages of EFL with the VOT patterns produced in English as an L3 by bilingual speakers of Pomeranian (L1) and BP (L2) (between group analysis).

The participants in this study were 40 children aged 8 to 10 years old studying in the 3rd grade of a public school. They were divided in two groups, one with 20 monolingual children, and another one with 20 bilingual children. The languages spoken by the monolingual children was Brazilian Portuguese; the bilinguals, in turn, spoke Pomeranian (L1), German (L1) and BP (L2) on a daily basis. English was studied as an L2 by monolinguals and as an L3 by bi/multilinguals.

The data was collected with the aid of an adapted version of the instrument created by RINALDI (2008), which is a fairy tale called “The fairy tales have gone nuts” and a tic-tac-toe game. The story-telling sessions took place in three different days and were performed in three languages: Portuguese, English and Pomeranian, and involved the participants in the most playful way as possible. The stories in each language contained target words beginning voiceless obstruents /p, /t/ e /k/ followed either by /a/, /i/ or /u/. After each story-telling session, the participants were recorded individually in a studio while playing the tic-tac-toe game. One of their tasks while playing the game was to use the target words in carrier sentences which were the „spells“ the participants had to cast on the fairy-tale characters. With this procedure, recordings of the voiceless stops in the three languages spoken by the participants were made over a period of a month. In

Portuguese 840 tokens (7 target words X 3 repetitions X 40 participants) of /p/, 840 tokens for /t/ and 840 tokens for /k/ were recorded. In Pomerannian 300 (5 target words X 3 repetitions X 20 participants) tokens for /p/, 300 tokens for /t/ and 300 tokens for /k/ were analyzed. In English 360 tokens (3 target words X 3 repetition X 40 participants) for /p/, 360 tokens for /t/ and 360 tokens for /k/ were measured.

The VOT measures were performed with the software Praat, version 5.1.04. (Boersma & Weenik, 2003). No statistical treatment has been given to the data so far, although t-tests will be run to compare the VOT of English voiceless plosives produced by bilinguals and monolinguals. However, descriptive statistics indicate that the VOT patterns prevailing when the multilingual children spoke English (L3) is the one similar to those of their mother tongue (Pomeranian). Considering that in Pomeranian the voiceless plosives are aspirated, we expected that in the other languages spoken by the multilinguals would present some aspiration of these segments. Indeed, we observed that the bilinguals produced very high values of aspiration, even in Portuguese, in which the voiceless obstruents are not aspirated (/p/: 48ms, /t/: 60 ms and /k/: 70 ms). The VOT pattern values for Portuguese are: 12 ms /p/, 18 ms /t/ and 38 ms /k/. In English the following values were found: 40 ms /p/, 56 ms /t/ and 65 ms /k/. The values for the English native speakers are: 55 ms /p/, 70 ms /t/ and 80 ms /k/ (Kent &Read, 1992). In Pomeranian, the mean values found were: 54 ms /p/, 44 ms /t/ and 76 /k/.

While in the literature the patterns of VOT for Brazilian Portuguese are 12ms for /p/, 18 ms for /t/, and 38ms for /k/, in this study the following means were found for monolingual children whose L1 is BP: 16ms for /p/, 17ms for /t/, and 40ms for /k/. The mean VOT values produced in English by the monolinguals were: 30ms for /p/, 38ms for /t/ and 52ms for /k/, while the mean values reported in the literature are 55 ms /p/, 70 ms /t/ and 80 ms /k/. The results, discussed within in a dynamic approach to multilingual speech production, highlight the gradient values in the intralinguistic transfer of VOT among the languages spoken by multilinguals.

References

Boersma, P. & Weenik, D. (2003). *PRAAT: doing phonetics by computer.* Version 5.1.04 (programa computacional). Available from http://www.praat.org.

Kent, R. & Read, C. (1992) *The acoustic analysis of speech*. San Diego: Singular Publishing.

RINALDI, L. (2008) *Procedimentos para avaliação instrumental da pronúncia do português brasileiro por crianças de 5 a 7 anos de idade.* CELSUL, Porto Alegre.

Near merger or lexical split? The case of mas and mais in Brazilian Portuguese

Michael Colley
michaeltcolley@yahoo.com

A potential case of a near merger is examined in Brazilian Portuguese, resulting from the diphthongization of vowels before word final /s/, such that mas 'but' and mais 'more' are both [maiʃ]. Although both vowels are fully diphthongal in Rio, /a/ in this case tends to be more fronted and raised than /ai/. Similar to near-mergers, some speakers make a distinction in their speech that they are unable to perceive. However, upon closer examination, the difference in production appears to be a lexical split driven by word frequency, rather than a near merger. All of the vowels examined in the pre-/s/ environment are becoming fronted and raised, but higher frequency words are more affected than lower frequency words. Thus the lack of symmetry between production and perceptual is not just a property of near-mergers, but may be a larger phenomenon.

Towards a Dynamic Computational Model of Portuguese Liquid Consonants

Michael Proctor.
mproctor@usc.edu

Many aspects of Portuguese phonology remain poorly described within feature-based frameworks, in part because it is not well understood to what extent phonological features are grounded in the phonetic domain. The liquid consonants /l/-/ʎ/-/ɾ/-/ʀ/ in particular, resist characterization in terms of distinctive features because of the great variation in their acoustic and articulatory realization. It is unclear why such a diverse group of sounds patterns together in the phonological organization of Portuguese and other languages (Walsh-Dickey 1997); neither has it been established what phonetic properties the members of the class might share (Ladefoged & Maddieson 1996). Better knowledge of the phonetic characterization of this group of sounds is therefore critical to our understanding of the nature of phonological representation.

In this paper, I introduce a computational model of Portuguese consonant phonology based on articulatory primitives (Browman & Goldstein 1992) being developed as an extension of the Task Dynamics Application (*TaDA*: Nam et al. 2004). Under this model, lexical items are represented as molecules of coordinated gestures. Hypotheses about phonological representations may be tested by comparing the results of articulatory and acoustic simulations against phonetic data acquired from Portuguese speakers producing the same utterances. Most importantly, by modeling liquid consonants in a gestural framework, we can appeal to general principles of dynamic organization of task-directed primitives (Saltzman & Munhall 1989) to account for their phonological behavior. Three phenomena of particular interest in Portuguese phonology which are being examined using this methodology are rhotic variation, coda liquid neutralization, and gestural organization in stop-liquid clusters.

A variety of features have been proposed to describe the shared phonological behavior of liquid consonants, including [lateral], [trill], [sonorant] and [liquid] (Walsh-Dickey 1997; Mateus & Pardal 2000), yet none of these primitives has proven adequate to capture the members of the class or their behavior in Portuguese (Mielke 2008; Proctor 2009). Crucially, feature-based models do not offer a principled account of the dynamic organization of phonological primitives.

The alternative hypothesis being examined in our work is that liquid consonants are

prototypically characterized as segments produced through the coordination of tongue-tip and tongue-body gestures (Giles & Moll 1975; Browman & Goldstein 1995). TaDA simulations based on these representations have so far been used to model a variety of phenomena relating to liquid consonants in Romance, Germanic, Dravidian and Slavic. The Spanish /ɾ-r/ contrast has been modeled as primarily resulting from differences in tongue tip stiffness. Russian /rʲ-r/ and /lʲl/ contrasts have been modeled as primarily resulting from differences in the specification of the dorsal constriction location. Phonological phenomena associated with post-vocalic liquids, such as coda deletion (Quilis 1999) and coloring and lengthening of nuclei (Wiese 2001) have been described as the result of blending of co-produced tongue body gestures.

Portuguese presents a special challenge for accounts of liquid representation because of the great variety of rhotics used in different varieties and by different speakers (Cruz-Ferreira 1995).

Varieties of Portuguese which use uvular rhotics are especially problematic for feature-based accounts of rhoticity. In the gestural model being developed here, each of the liquid consonants of Portuguese are characterized by the presence of an intrinsic tongue body constriction. These consonants differ in their specifications for constriction location and degree, and the way in which the tongue body is coordinated with tongue-tip gestures. Rhotic variation – both synchronic and diachronic – is argued to result from changes in gestural constituency and coordination. Coda liquid neutralization is modeled as a change in the organization of gestural primitives with respect to syllabic nuclei. Constraints on segmental organization in Portuguese clusters are argued to result from issues relating to production and recoverability of concurrent gestures.

TaDA simulations are presented which capture many of the fundamental acoustic and articulatory properties of Portuguese liquids, and compared with preliminary phonetic data. These models are described in the context of a broader research program whose goal is the development of a formal model of a Portuguese articulatory phonology.

References

Browman, C. & L. Goldstein. 1992. *Articulatory phonology: an overview.* Phonetica 49: 155-180

Browman, C. & L. Goldstein. 1995. *Gestural syllable position effects in American English*. In Bell-Berti & Raphael (eds.), Producing speech: contemporary issues, 19-34. New York: AIP Press.

Cruz-Ferreira, M. 1995. *European Portuguese*. JIPA 25(02): 90–94

Giles, S. & K. Moll. 1975. *Cinefluorographic study of selected allophones of English /l/*. Phonetica 31: 206–227

Mateus, M.H. & d'Andrade, E. 2000. The Phonology of Portuguese. Oxford: Oxford University Press.

Mielke, J. 2008. The emergence of distinctive features. Oxford: Oxford University Press.

Nam, H., L. Goldstein, E. Saltzman & D. Byrd. 2004. *TADA: An enhanced, portable Task Dynamics model in MATLAB*. JASA 115: 2430

Proctor, M. 2009. Gestural Characterization of a Phonological Class: the Liquids. Ph.D. Dissertation, Yale University.

Quilis, A. 1999. Tratado de fonología y fonética españolas. Madrid: Gredos.

Saltzman, E. & K. Munhall. 1989. *A dynamical approach to gestural patterning in speech production*. Ecological Psychology 1: 333-382

Walsh-Dickey, L. 1997. The Phonology of Liquids. Ph.D. Dissertation, UMass Amherst.

Wiese, R. 2001. *The phonology of /r/*. In T.A. Hall (ed.), Distinctive feature theory, 335-368. Berlin: Mouton de Gruyter.

Factors of disagreement in the speech segmentation of the C-ORAL-BRASIL corpus

Priscila Côrtes.
pritico@gmail.com

This study derives from the project C-ORAL-BRASIL, which focuses on the compilation of a corpus of spontaneous spoken Brazilian Portuguese (200 texts and 300,000 words divided in formal and private monologues, dialogues and conversations of informal and formal speech) according to the criteria used by its European project counterpart, the C-ORAL-ROM, responsible for the compilation of corpora of the four major Romance languages (French, Italian, Portuguese, Spanish). The corpora are designed to allow for the study of illocutions (AUSTIN, 1962). The Brazilian project aims to study Brazilian Portuguese comparatively to the languages that compose the European project. Both projects take into account the "Teoria della lingua in ato" Speech Patterning Theory (CRESTI, 2000), which considers prosody as the interface between the linguistic domain (the utterance) and the action domain (the illocution). The utterance is the shortest linguistic unit that can be pragmatically interpreted, to which there is necessarily an illocution associated. Speech is segmented into utterances and tonal units according to intonational perceptual breaks. A conclusive break (//), perceived as terminal, determines the end of an utterance. Inconclusive breaks (/), perceived as non-terminal, mark the end of a tonal unit and therefore the beginning of another.

The students who are associated to the project have been through a thourought training on segmentation, which included workshops, courses and tests. The students were separated at first in three groups, Group 1 consisting of the most able three segmenters, Group 2 consisting of the second best three segmenters and Group 3, with four members, who were novice to the project at the time tests were undertaken. However, as some project members discontinued their training, groups 2 and 3 were merged and into a single group, named Group 2, with four members.

The tests undertaken by the groups consisted of a segmentation of a certain text, done individually by the members of the groups. The segmented texts were submitted to kappa tests and the results were discussed by the students members of the project and one of the coordinators. During the segmentation process, students were to segment the transcription basing themselves only on their perception of the breaks.

Due to the perceptual character of the process, and despite the training, there are irregularities in the results, since there are situations in which the transcribers do not have the same segmentation patterns - for example, one considers the break as terminal whereas the other, as non-terminal, or two don't notice the break, whereas the other marks a non-terminal, or even a terminal break. Nevertheless, it is possible to perceive contexts that are more likely to generate different decisions from the transcribers, thus the discrepancy in the results. The comparison of the results of the kappa test has shown progression in Group 2 results from the beginning of the process; Group 1 seemed to be sufficiently trained. This study aims at finding the factors behind the disagreements between transcribers during the training process.

These factors vary from transcribers' level of attention to induction generated by the theory in which the segmentation itself is based upon. Furthermore, disagreements have a recurrent pattern - theory induction happened in spite of the first instruction given to the groups, that is, that they should only base themselves in their own perception.

This study consists of an analysis of the 7 transcriptions segmented by the groups during the training process and submitted to kappa tests. So far the reasons of disagreement can be separated in the six following major groups: phonetic factors (speed of the speech, coarticulation, overlapping), prosodic factors (accents, variations of pitch, ascendant and descendant endings, rhythm, emphasis), syntactic factors (theory induction, syntactic influence), semantic factors (influence of the semantic content of the unit or utterance), cognitive factors (distraction and marking mishaps) and factors related to the transcription (changes in the transcription criteria, non-identified word and incorrectly transcribed fragment). The factors distraction of the transcribers and problems in the transcription, which were of average quality, seem to rule the results. There is also great discrepancy around interrupted utterances and retractings (problems of execution, i.e., repetition of a segment with a similar semantic value to the one that has been canceled), probably because most of the transcribers didn't have, of course, much experience in segmenting and therefore couldn't tell cases of interruption from retractings. Another focus of disagreement is ascendant endings of some utterances, normally associated with a non-conclusive profile.

A more detailed analysis of the data will show the differences between the 2 groups, which could be associated with different factors, for example, would Group 1, which has had more contact so far with the theory, be more inducted by it? On the other hand, would Groups 2 and 3 be more semantic conditioned, since they have had less contact with the theory and less time to practice segmentation?

The results of this study are relevant to the members of the project, because once the contexts which are likely to generate disagreements in judgement are identified, the quality of the

segmentations to be done can be raised, because the researchers would be more aware of those contexts, therefore being more attentive to them.

Speech Production of mild Alzheimer's disease patients

Renne Alegria and Maria Ines Nogueira.
realegrias@usp.br

Background:

Little attention has been given to the population with Alzheimer's disease concerning the sociolinguistic underpinnings of their oral discourses and speech productions. Language performance research has been the focus recently, especially for advancing our understanding of the multiple cognitive and speech production on the patients with dementia. Objective: to demonstrate the preservation of speech production of Alzheimer's disease patients.

Methods

The participants for this study were eight Alzheimer's disease patients and six healthy controls. The patients were four men and four women aged 80 and older, and the controls were three male and three female, aged 75 and older. All the patients were from PROTER- Old Age Program - ambulatory care of the Institute of Psychiatry of the School of Medicine of the University of São Paulo. The controls were their caregivers and/or relatives. The patients and controls had 4 to 11 years of education, their Mini Mental State Exam (MMSE) scores 13 to 30. The controls were also evaluated by other scales such as SRQ-20.Their conversations were recorded and analyzed by the computational tool Stablex, based on mathematical-statistical-computer assisted program which mainly distinguish the results of preferential, basic and differential vocabulary.

Results

The speech production on the discourses of the patients is not necessarily destroyed compared to the discourse of the control subjects. The program Stablex shows that preferential vocabulary is constituted of the thematic vocabulary and the lexical components of more sociolinguistic frequency in the discourses. The patients produced more preferential words than the controls and they were shown to be statistically significant.

Conclusions

Assessing the discourse performance of patients with Alzheimer's disease provides opportunities to unravel the connections and changes among language, communication and behavior. For patients with dementia, analysis of speech is an access for a better communication with their caregivers

Emergent Consonantal Clusters

Ricardo F. Napoleão.
ricardonapoleao82@hotmail.com

This paper aims to discuss the emergence of consonantal clusters in Brazilian Portuguese (henceforth BP). It will be shown that emergent consonantal clusters follow from the deletion of an unstressed vowel which is adjacent to a sibilant. The BP data to be discussed come from the state of Minas Geris, Brazil and Exemplar models will be adopted to theoretically support the analysis (Johnson 1997, Pierrehumbert, 2001).

Traditionally, the syllable structure of BP has been described as being predominantly of a CV type, although CVC and CCV syllables also occur in the language (CAMARA JR, 1985; COLLISCHONN, 1996, CRISTOFARO-SILVA, 1996). Consonantal clusters tend to undergo phonological processes in the variety of Minas Gerais. Thus, CCV syllables are reduced to CV ones as in li[vr]o > li[v]o “book” and CVC syllables are also reduced: fe[st]a > fe[s]a “party” (CRISTOFARO-SILVA, 1999). These two cases of cluster reduction indicate that phonological processes seem to favor CV syllables. Another group of consonantal clusters involve typically two obstruents, as in a[ft]a “mouth ulcer” and pa[kt]o “pact”, and these clusters, unlikely the previously presented ones, are rather separated by an epenthetic vowel: a[fit]a and pa[kit]o.

Although in this case epenthesis is favored over cluster reduction one observes that once again a CV syllable seem to be the end result. Thus, the data presented so far seem to indicate that cluster reduction and epenthesis would favor CV syllables, which are already predominant in the language.

One would expect that BP would be moving towards a consistent CV syllable type.

Nevertheless, phenomena related to vowel deletion have created unexpected consonantal clusters that had not been previously displayed in the language. For example, he[hps] “herpes” has a more conservative pronunciation as he[hpis]. This phenomenon is also observed word initially as in [dst]ante “distant” which has a conservative pronunciation as [dist]ante and also observed word internally as in pa[htsp]ação “participation” which may also be pronounced as pa[htisip]ação. All cases involving consonantal clusters we have presented as emergent patterns show that a front high vowel [i] may break up the consonantal cluster.

Besides a high front vowel one notices that emergent consonantal clusters may involve a

high back vowel. Some examples are: word initially as in [k^{w}st]ura ~ [kust]ura “s/he sews”, word finally as in bar[k^{w}s] ~bar[kus] “boats” or word internally as in com[p^{w}st]ura ~ com[pust]ura “composure”.

The emergence of consonantal clusters in BP has been dealt with in the literature from a phonological perspective (BISOL; HORA, 1993; ABAURRE; PAGOTTO, 2002 *apud* LEITE, 2007). Experimental work has involved mainly cases related to a high front vowel (ALBANO, 2001). In an attempt to expand previous findings and further investigate these phenomena, we have designed an experiment to evaluate vowel reduction and the emergence of consonantal clusters.

All emergent consonantal clusters which have been presented in this paper involve the a high vowel adjacent to a sibilant which lead us positing the hypothesis that the high vowel is in fact reduced and temporally merges with the sibilant. To test this hypothesis we analyzed cases where the vowel is or is not present as in he[hps] ~he[hpis] “herpes”. All the words were tested in similar prosodic environments. Participants were university students who were not aware of the aim of the project. Recordings had good quality in order to support the experimental analysis.

The acoustic analysis shows that although a full vowel cannot be identified as being adjacent to the sibilant it is possible to observe that vowel deletion leaves relevant traces in the signal. It is argued that the fine phonetic detail which can be observed in the acoustic analysis of emergent consonantal clusters in BP actually play an important role in language change. It is suggested that Exemplar Models can accommodate these findings by clustering exemplar which are related. Further work may consider token and type frequency effects in the analysis of the phenomenon. Perceptual experiments may be run to understand better the relationship between perception and production.

References:

ALBANO, Eleonora Cavalcante. *O gesto e suas bordas:* esboço de fonologia acústicoarticulatória do português brasileiro. Campinas: Mercado das Letras, 2001. 272 p. (Leituras no Brasil)

CÂMARA JR, Joaquim M. *História e estrutura da língua portuguesa*. Rio de Janeiro: Padrão, 1985. 256 p.

COLLISCHONN, G. A sílaba em português. In: BISOL, L. (Org.) *Introdução a estudos de fonologia do português brasileiro.* 3. ed. rev. Porto Alegre: EDIPUCRS, 2001. Cap 2, p. 91-122.

CRISTÓFARO-SILVA, Thaïs . Sobre a Quebra de Encontros Consonantais no Português Brasileiro. Estudos Lingüísticos (São Paulo), São Paulo, v. 29, p. 522-527, 2000.

_____. *Fonética e fonologia do português: roteiro de estudos e guia de exercícios*. 2. ed. São Paulo: Contexto, 1999. 254 p.

LEITE, Camila T. *Sequências de oclusiva alveolar + sibilante alveolar como um padrão inovador no português de Belo Horizonte*. 2007, 126f. Dissertação (Mestrado em Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, 2007.

CV phonotactical biases in Brazilian Portuguese acquisition

Rosane Garcia and Márcia Zimmer.
garcia.rosane@gmail.com

A number of studies investigate the role of biomechanics in shaping CV co-occurrence biases found in babbling and first words in many languages (e.g. MacNeilage & Davies, 2000). Based on the articulation of stop (/p, b, t, d, k, g/) and nasal (/m, n, N/) consonants with certain vowels, the universal CV biases are: 1) LC - labial (L) C's tend to occur with central (C) V's (LC) ; 2) CrF - coronal (Cr) C's favor front (F) V's; 3) VB - velar (V) Cs favor back (B) V's; Such findings support the Frame then Content theory, which predicts that inertia is the biomechanical basis of canonical syllable, and the regular mandible movement (frame) would precede the richer linguistic phonic content which would emerge later.

The aim of this study, which is part of a doctoral research in progress, is to investigate whether the CV combinations described in the literature are observed in data acquisition concerning the southern Brazilian Portuguese variant. Furthermore, we try to avoid some methodological flaws committed by MacNeilage & Davis (2000) regarding their statistics analysis, which is limited to observed-to-expected (O/E) ratios derived from chi-squared tables, without any effect size measure.

The two corpora of this study – the children corpus and the adults corpus – were derived from 100 hours of recordings of spontaneous speech during child-adult interactions.

Ten children, aged 1:0 to 3:0 years old, were recorded during spontaneous talk with their caregivers for approximately an hour every 30 days during an average period of a year. All recordings were orthographically transcribed and translated into phones with Ortofon (ALBANO & MOREIRA, 1996). Frequency counts of around 20,000 items were performed with WordSmith Tools (SCOTT, 2008). The analysis was divided into two age-groups: from 12 to 23 months, and from 24 to 36 months. The CV biases counts include the 19 consonants of Brazilian Portuguese: the labial /p, b, v, f, m/, the coronal /t, d, s, z, ʃ, ʒ, l, ʎ, r, n, nh/ and the dorsal /k, g, R/ occurring with the front, center and back vowels. Statistical analysis of CV biases included chi-square (χ^2), Cramer's V (φ) and Freeman-Tukey deviates (FT).

The analysis between the two databases (children and adults corpora) indicates association between them with X^2= 329,02, p = 0,001, φ= 0,10. In the first phase of analysis concerning the data of 12 to 23 month-old children, the CV biases reported in the literature were confirmed: biases

indicate propensity for production of L with F (FT 2.69), Cr with F (FT 4.55) and D with B (FT 5:48). However, Cramer's V coefficient (0.05) indicates a low statistical significance between pairs. At this phase of analysis, the patterns observed in adult speech differ with respect to the labial (L) consonants, indicating preference with F (FT 2.48), but the biases Cr F (FT 3, 06) and DB (FT 6.88) were in line with the preference reported in the literature.

In the second phase of analysis, the associations provided for child speech (X^2= 119; φ= 0,09) are the same as the ones observed in adult speech data (X^2= 160; φ= 0,06). CV cooccurence biases in infant and adult speech, respectively, coincided: LF (FT 2,94; 3,16), Cr F (FT 2,14; 2,30) and DB (FT 4,61; 6,96). Table 1 shows the CV biases according to the two stages of analysis of the two groups. Values in bold identify the results differing significantly from those expected.

Preliminary results partially confirmed the findings of MacNeilage and Davis. However, the estimates concerning the CV biases are unstable in the acquisition data of Brazilian Portuguese, because of the low statistical power indicated by the analysis. Although our corpora have a bigger size than those used by MacNeilage and Davis (2000), the results call for the need to increase the number of types, which is being arranged with the transcription of recently collected data. Moreover, further analysis is needed of the combinations of biases in other syllabic contexts within the words (e.g. unstressed initial position, medial stressed position, unstressed final position), which would take into consideration prosodic and grammatical factors. Albano (2009), in a study performed with greater corpora of Brazilian Portuguese and Spanish lexicons of adult data, reports that CV biases were found only in unstressed initial position.

Finally, given the remarkable emphasis placed on the speech dynamics characterizing phonological processes (Browman & Goldstein, 1992), a study on CV biases could not center its explanation on the biomechanical inertia. Therefore, we spouse the view forwarded by Gestural Phonology, according to which it is the coordination among the gestures of articulators – which are potentially independent from the beginning of the process of language acquisition – that plays a central role in acquiring the phonic inventory.

Table 1 - CV biases found in children and adult corpora in two stages of analysis

Cplace	Vplace	Phase 1 (12 – 23 months)						Phase II (24 – 36 months)					
		Group 1			Group 2			Group 1			Group 2		
		F	**C**	**B**	**F**	**C**	**B**	**F**	**C**	**B**	**F**	**C**	**B**
Labial	Observed	240	370	234	755	547	507	691	480	461	770	513	485
	Expected	269	320	254	688	538	582	616	512	504	685	537	545
	FT Dev.	-1,78	**2,69**	-1,32	**2,48**	0,38	-3,21	**2,94**	-1,42	-1,95	**3,16**	-1,04	-2,66
Coronal	Observed	442	382	280	1785	1244	1328	1389	1055	1027	1814	1336	1283
	Expected	352	418	333	1658	1297	1402	1310	1088	1071	1717	1347	1368
	FT Dev.	**4,55**	-1,82	-3,04	**3,06**	-1,46	-2,00	**2,14**	-1,02	-1,37	**2,30**	-0,29	-2,34
Dorsal	Observed	97	176	498	281	415	551	239	392	409	286	402	519
	Expected	158	188	150	474	371	401	392	326	321	467	366	372
	FT Dev.	-5,45	-0,93	**5,48**	-10,02	2,22	**6,88**	-8,68	3,48	**4,61**	-9,40	1,80	**6,96**

Reference:

ALBANO, E.C. (2009). *A corpus-based reappraisal of the role of biomechanics in lexical phonotactics*. Corpus Linguistics 2009 Abstracts. Liverpool: University of Liverpool, v. 1. p. 48.

BROWMAN, C.; GOLDSTEIN, L. (1992). Articulatory Phonology: an overview. *Phonetica*, 49, 155-180.

MACNEILAGE, P.F.; DAVIS, B.L. (2000). On the origin of internal structure of word forms. *Science*, 2000, 527-531.

Comparing the 'writing production', 'perception' and 'oral production' tripod in monolingual and bilingual data

Sabrina Gewehr-Borella, Márcia Zimmer and Ubiratã Alves.
binagewehr@yahoo.com.br

The number of research studies that investigate the relationship between perception and oral production has increased in the last years. In spite of the significance of this kind of studies, which relate oral perception and production, we believe that is also necessary to include the study of another aspect: writing production. For this reason, in this study we compare Voicing pattern changes in plosive segments, as we consider data, collected from monolingual and bilingual students, from three different sources: writing production, oral production and perception.

Our participants were divided in three groups: a) Monolinguals, who speak only Brazilian Portuguese and do not have contact with bilingual speakers (group M1- 87 participants from Rio Grandecity); b) Monolinguals, who speak only Brazilian Portuguese and do have contact with bilingual speakers (group M2- 30 participants (from Picada Café/RS town)); and c) Bilinguals, speakers of Brazilian Portuguese and of the German Immigration Language called Hunsrückisch (group B -66 participants (from Picada Café/RS town)).

The research study was divided in three phases. In phase 'A' (writing production), we analyzed, firstly, the number of grapheme changes 'p,b', 't,d' and 'c,g' by 183 students (*general analysis*) from preschool to the fifth grade of Elementary School, from the three participating groups. Secondly, we analyzed, separately, the written data of 30 students (10 from each group-*partial analysis*) from the 183 previously participants. Written data were collected through a game (bingo), in which each participant should write 18 words initiated by the graphemes 'p,b,t,d,c,g' (3 of each one). In phase 'B' (perception), we analyzed the results of a perception test, in which the 30-student-group took part. The test, consisted of the perception of 18 words presented within a carrier sentence. The 18 sentences were divided in the following way: a) 6 words initiated by a voiceless plosive; b) 6 words initiated by a zero VOT plosive, manipulated by us from a voiceless plosive; and c) 6 words initiated by a voiced plosive. The test was carried out in the following way: After hearing the sentence "A palavra é…cola (The word is…glue)", the student needed to check: option 'a' (a cartoon of a glue), if he heard the Brazilian Portuguese word [kçla] (glue); option 'b' (a carton of a question mark), if he was in doubt; and option 'c' (cartoon of a collar), if he heard the

word [gçla] (collar). With this test we analyzed, firstly, the number of mistakes made by the participants from the three groups, in the questions whose VOT values had not been manipulated. After this, we verified, in the manipulated options, which option the participants preferred to check (more for the voiceless or for the voiced words), while hearing a word initiated by a zero VOT. Phase 'C' (oral production) was divided in two moments. Firstly, acoustic analyses measuring VOT values in Brazilian Portuguese in the plosives produced by 30 participants were carried out with *Praat* (*version 5.0.3.2*). After this, we carried out acoustic analyses of VOT in the Hunsrückisch data, which had been produced by 10 out of 30 students, who participated in the three phases of the research. For the collection of the target words, participants were shown PowerPoint slides with drawings of the target nouns initiated by plosive segments, which should be spoken in Portuguese Within the carrier sentence " Fale (palavra-alvo) de novo (Say (target word) again)" and in Hunsrückisch in the carrier sentence "Só (palavra-alvo) kanz xen", which means, "Say (target word) very beautifully".

For the analyses, of the three phases, were apllied Kruskal-Wallis and Wilcoxon nonparametric statistic tests. With regard to the results obtained, in phase 'A' (writing), we found more graphemic changes in the bilingual students (group B), followed by the monolingual students from group M2 and then from M1. As we can see, the order presented is related to bilingualism. Although the means follow the same order in the *general analysis* and in the *partial analysis*, significant differences in the changes (p=0,0001) were found only in the *general analysis*. In phase 'B', in the non-manipulated words, more perception mistakes were found among members of the M2 group, followed by group M1 and then by group B. In this analysis, significant differences were not found (p=0,5315). In the manipulated words, group M1 checked option 'c' more times (voiced), without a significant difference (p=0,0769) between this option and option 'a', group M2 chose option 'a' more frequently (voiceless), without showing significant difference (p=0,1164) with option 'c' and group B chose option 'a' more times (voiceless), with significant difference (p=0,0301). In phase 'C', part of the measurements shown in Portuguese VOT values matched our expectations, as the M1 group exhibited the shortest VOT, whereas the bilingual group showed the longest VOT values. In some negative VOT measurements, M1 exhibited the longest pre-voicing values, whereas the bilingual group showed the shortest pre-voicing values. The only significant difference found occurred in the plosive segment [k] (p=0,0273). In the analysis of Hunsrückisch words, as hypothesized , Only zero and positive VOT values were found in word-initial position.

Significant differences were not found in any of the groups analyzed when comparing the three phases of the study (writing production, perception and oral production).

The analysis of the data revealed two kinds of transfer: 1st) Grapho-phonicphonological

transfer (ZIMMER, ALVES, 2006): verified in the comparisons of writing and oral production, as for instance, in the production of [»pç.tSi] for the word ‘bote’ (boat) and in the writing of ‘pola’ for the word ‘bola’ (ball); 2nd) Phonetic-phonological transfers (FLEGE, 2002; BEST, TYLER, 2007): linked to the dynamic aspect between the physic phone and the phoneme (ALBANO, 2001, 2007) and generated in reason of the perception difficulty of the acoustic-articulatory differences between the L1 and the L2.

Our data provide additional evidence for the dynamic view of phonology which guides this study, as we notice gradient values in the phonetic-phonological transfers analyzed. We therefore conclude that some of our participants show a positive correlation rate among the grapheme changes ‘p,b’, ‘t,d’, ‘c,g’ and the perception and speech production errors that correspond to these segments, which suggests a possible relation between the processes of writing production, oral production and perception.

Phonetic characterization and socio-linguistic stratification of the vowel allophones of the Spanish of Concepción (Chile)

Scott Sadowsky
ssadowsky@gmail.com

There is near-universal consensus that the vowel system of Spanish consists of five phonemes (/i e a o u/), each having a single allophone, typically defined as [i e a o u]. Additionally, the existence of assimilation-based nasalized allophones ([ĩ ẽ ã õ ũ]) in the presence of nasal consonants is commonly recognized.

An exploratory study of the Spanish spoken in Concepción, Chile, however, reveals a previously unattested diversity in vowel allophony along both the F1 and F2 dimensions.

Furthermore, these allophones appear to be socio-linguistically conditioned.

This paper presents the preliminary findings of the exploratory study, which analyzed the vowels of 16 female informants of between 16 and 22 years of age, all monolingual Spanish speakers and life-long inhabitants of Concepción, Chile. Informants' monthly family income ranges from US$ 280 to US$ 3600, while the maximum educational level achieved in their nuclear family groups ranges from vocational high school to an undergraduate university degree. 800 stressed vowel tokens occurring in the context of voiced and voiceless stops and fricatives were elicited by means of specially-designed sentences and a wordlist embedded in carrier phrases, and were recorded using a Studio Projects C-1 large-diaphragm condenser microphone, a Studio Projects VTB-1 pre-amplifier, and an M-Audio FastTrack Pro audio interface. The tokens were analyzed using the Praat software package by selecting the central third of each vowel nucleus and calculating their average F1 and F2 values using the Burg algorithm. The resulting values were then normalized using the Nearey 2 algorithm.

The analysis shows that normalized vowel allophones vary widely between speakers.

Using the quasi-cardinal vowels proposed by Lindblom (1986) as a point of reference, the various allophones observed for each vowel phoneme occur in an acoustic space bounded approximately as follows: /i/ → [i̠ ɪ̝ e̝], /e/ → [e̞ e̞̠ ɛ̽ ɛ], /a/ → [ɐ̝ ä a̠], /o/ → [o̹ ɔ ɔ̇] and /u/ → [u̠ u o ʊ]. As stress, phonetic environment and linguistic style were controlled, the distribution of these vowel allophones can be attributed to socio-linguistic factors. Indeed, an analysis of the vowel allophones by socio-economic level shows a pattern of stratification with a clear distinction

between each of the four groups studied: Upper Middle → [ɪ ̠ ɛ ̠̽ ɐ ˔ ʊ̞ u̠], Middle → [i ̠ ɛ ̟̽ ɐ ʊ̞ u̠], Lower Middle → [ɪ ɛ̽ ɐ ʊ̞ u̠], Lower → [ɪ ɛ̽ ä ɔ̽ u̞].

Given that the existence of multiple allophones of vowel phonemes has not been previously attested in Spanish, this phenomenon appears to represent a case of linguistic change, possibly on-going. These results call for further investigation of the acoustic and socio-linguistic aspects of the vowel system of Chilean Spanish, and possibly other national varieties of Spanish as well.

Prosodic constituent structure and the spread of anticipatory pharyngealisation in Arabic Libyan Arabic

Tareq Maiteq
tareqmaiteq@ling.ed.ac.uk
Department of Linguistics and English Language, University of Edinburgh, UK

The hierarchical effect of prosodic constituent structure on co-articulation has implications for models of speech production (Cho 2004, McClean 1973, Kent & Minifie 1977). We explore such effects by looking at the production of emphatic (pharyngealised / T, D, S/) consonants in two Arabic dialects, i.e., Libyan and Egyptian. We seek to quantify how the degree and extent of anticipatory pharyngealisation, induced by these pharyngealised consonants, vary as a function of prosodic boundary level (syllable vs. word vs. phrase vs. intonation phrase). It is hypothesized that vowels at higher prosodic boundaries are co-articulated less with adjacent (post-boundary emphatics) than when at lower boundaries.

Our carefully controlled data were constructed as follows:

[V1bV2 + boundary type + post-boundary emphatic T, D, S].

Syllable: [m**abaa.D**iʃ]

Word: [dab**aaba**#**D**aayʕa]

Phrase: [ibn ʕam idab**aaba+D**aaʕ]

IP: [ħasb kalaam iddab**aaba%D**aaʕ]

Test vowels include all possible combinations of [a] and [i], i.e., [aba; abi; iba; ibi]. Although the number of items varied according to condition, we included an average of 10 items per condition, and had 2 repetitions for six speakers of each dialect. Segmentation criteria we followed here were based on defining constriction onsets and releases as presented in Turk et al (2006). First-pass segmentation was performed on zoomed out spectrogram displays for defining general boundary regions (within an accuracy of 5-10 ms). A fine-grained segmentation procedure was then performed by more zoomed in waveform displays.

In all our /VbV/ test sequence segmentation, stop closure onsets were defined at F2 offset, which coincide with an overall dip in amplitude. Stop releases were located at the first release burst. When the burst of the stop release was not clear then it was be marked near the point of F2 onset. The VOT interval was included as part of the duration of the vowel. This is so to make reliable

comparability of [T], for example, with [S] and [D] (voiceless stop, fricative and voiced stop, respectively). Fricative boundaries, e.g. [S] were marked at onsets/offsets of frication energy that acoustically correlate to sibilant place of articulation. In pre-pause positions, the segmentation criterion for vowels is based on continuous F2 energy. The pharyngealised segments were marked at their release in post-pause positions, i.e., in the intonation phrase data set. This is because the constriction release in this case in difficult to determine.

F2 measurements at [V1 and V2] intervals were obtained at:

a) Absolute points (every 10 ms backwards from the consonant constriction onset),

b) Three proportional points for each vowel interval.

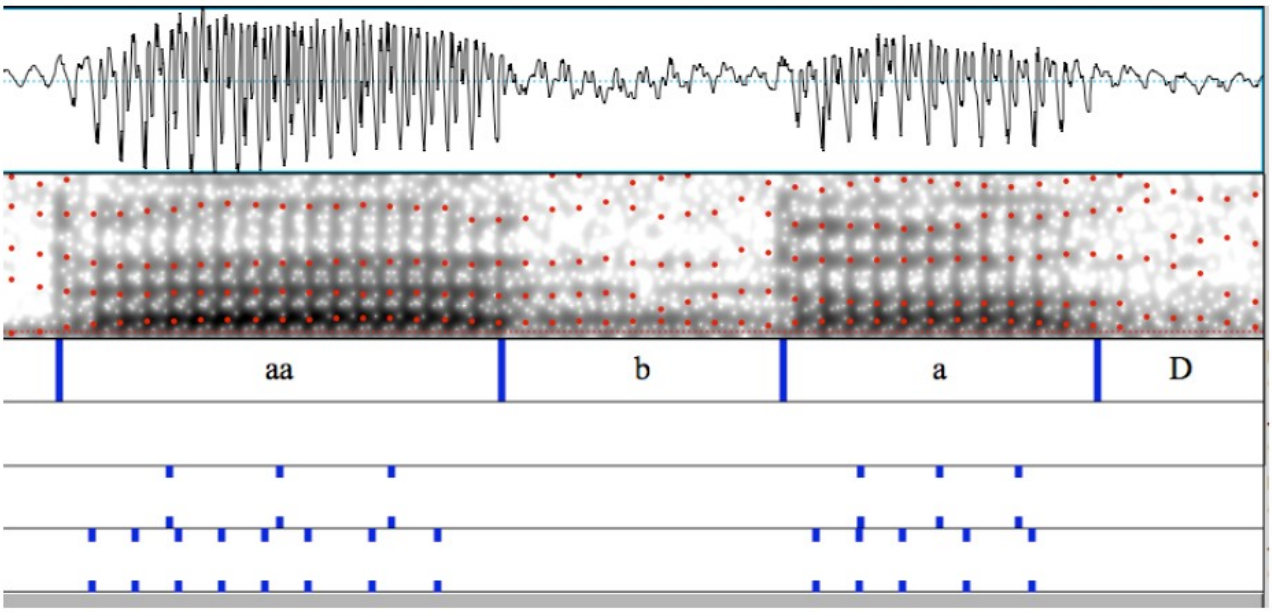


Fig (1): a spectrogram showing measurement points, a) Tier 3: three proportional points (b) Tier 4: absolute points marked at every 10 ms backward from the trigger.

The rationale underlying these measurements is to provide quantitative data about the degree and extent of pharyngealisation in each boundary condition. Our results will be interpreted in terms of phonological feature spreading vs. coarticulation accounts.

References

Cho, T. (2004). Prosodically conditioned vowel-to-vowel coarticulation in English. Journal of Phonetics, 32, 141-176.

Kent, R. D. and F. D. Minifie (1977): Coarticulation in recent speech production models. Journal of Phonetics 5: 115-133.

McClean, M. (1973): Forward coarticulation of velar movement at marked Junctural Boundaries, Journal of Speech and Hearing 16: 286-96.

Turk, A. E., Nakai, S. and Sugahara, M. (2006): Acoustic segments duration in prosodic research: A practical guide, in Sudhoff, S.,Lenertova, D., Meyer, R., pappert, S., Augurzky, P., Mleinek, I., Richter, N., and schlieber, J. (eds), Methods In Emperical Prosody Research, Walter de Gruyter, Berlin, New York.

On the nature of raised mid front vowels in Brazilian Portuguese

Victor Hugo M. Soares and Liliane P. Barbosa.
viktormedina@gmail.com

This research reports on the nature of raised mid front vowels in Brazilian Portuguese (henceforth BP). Mid vowels, both front and back, might be raised in BP as the following examples show: p[e]rigo > p[i]rigo „danger‟ or c[o]ruja > c[u]ruja „owl‟. A great amount of research has been carried out to investigate the raising of pretonic mid vowels in different dialects (cf. Bisol, 1991; Bortoni *et al.,* 1992; Oliveira, 1991, 1992; Viegas, 1987).

In this paper we will focus on the investigation of raised mid front vowels to high vowels and evaluate our findings in contrast to other front high vowels that occur in BP. We will also focus our investigation on vowels occurring in pretonic position. This is due to the fact that pretonic vowels are better defined acoustically than postonic vowels, since the latter are typically reduced. There are three different instances of pretonic high front vowels [i] in Brazilian Portuguese (PB): i) regular vowel, systematically realized as [i], as in b[i]chinho “pet”; ii) epenthetic vowel, optionally absent as in ad[i]stringente ~ ad[]stringente “binding” and iii) raised vowel, optionally realized as either [i] or [e], as in p[e]rigo ~ p[i]rigo “danger”. The cases listed in (i) and (ii) were investigated by Cristófaro-Silva and Almeida (2008). The authors found that regular and raised mid vowels are present in lexical representations because their presence in a word is favored. Concerning the duration of those vowels, their study shows that regular high front vowels are significantly longer than the epenthetic vowels. In addition, it states that the epenthetic vowel is favored when a voiced consonant follows it. Since Cristófaro-Silva and Almeida (2008) found evidence for fine phonetic details differentiating regular vowels from epenthetic vowels it seems appropriate to posit the question whether raised vowels would present any specific difference from the regular and epenthetic vowels. If there is any difference, we should be able to identify it and try to characterize it in relation to the other two types of pretonic high vowels investigated by Cristófaro-Silva and Almeida (2008).

The theoretical basis for this study is the exemplar model (Pierrehumbert, 2001, 2003). This model relates lexicon and grammar, besides claiming that mental representations are multidimensional and rich. It states that they are connected in exemplar clouds which are

constituted of recurrent words and chunks of the language. These are, in their turn, stored in memory taking into consideration all the variations the speaker may be exposed to through speech perception and production. The clouds are organized and stored accordingly to their regularity and phonological and semantic similarity degree. Therefore, those representations are a reflex of the range of neighboring memories in the exemplar clouds. These mental representations are "gradually built up through experience with speech" (Pierrehumbert, 2001, p. 1). Furthermore, the same phonetic signal may be connected to more than one representation. For these reasons and because this study is based on mental representations of sound structures, as well as their physical representations, methodological procedures following Laboratory Phonology principles will be adopted.

The data for this study was collected from five women and five men aged between 20 and 25, all being university students. Participants were asked to read as spontaneously as possible a set of ten sentences in which the data to be investigated were contextualized sentences containing each tested word. A number of distractor sentences were added at random between the sentences containing the targeted sentences. Each participant repeated the sentences four times, which generated a total of 40 utterances.

The initial analysis consisted of auditory discrimination of each token and edition of individual files to be analyzed acoustically. Each file was then examined individually and for each token we measured it was analyzed the values of F1 and F2 and duration. The analysis was done on *Praat*. Results indicate that phonetic detail might be relevant in some cases but not in others. That means that lexical organization is relevant in the mapping of phonological information. One potential explanation for this result is that some pretonic raised vowels may assume similar configurations of regular high front vowels as a consequence of merging. If this is the case, we would posit the hypothesis that frequency effects would have a role to play in implementing the merging of the two vowels in question. Further research may consider other structural environments in relation to adjacent consonants as well as whether the vowel is in a strong or weak position of a foot.

References

BISOL, L. (1981). *Harmonização vocálica*: uma regra variável. PhD Thesis. UFRJ. Rio de Janeiro.

BORTONI, Stela M; GOMES, Christina A.; MALVAR, Elisabete (1992). A variação das vogais médias pretônicas no português de Brasília: um fenômeno neogramático ou de difusão lexical? *In: Revista de estudos da linguagem*. v. 1. UFMG. Belo Horizonte.

CARVALHO, Sarah de Araújo; PARLATO-OLIVEIRA, Erika; SILVA, Thaïs Cristófaro Alves

(2006). Análise experimental da vogal epentética do português brasileiro. *In: REUNIÃO ANUAL DA SBPC*, 58, 2006, Florianópolis. Anais eletrônicos... São Paulo: SBPC/UFSC, 2006. Disponível em: http://www.sbpcnet.org.br/livro/58ra. Em edição. ISBN: 978-85-86957-11-6.

CRISTÓFARO-SILVA, Thaïs; ALMEIDA, Leonardo (2008). On the nature of epenthetic vowels. In: BISOL, Leda; BRESCANCINI, Claudia R. (ed.). Contemporary phonology in Brazil. Cambridge: Cambridge Scholars Publishing.

OLIVEIRA, M. A. (1991). The neogrammarian controversy revisited. *In: International journal of the sociology of language*. Berlin. v. 89, 93-105.

_____ (1992). Aspectos da difusão lexical. *In: Revista de estudos da linguagem*. ano 1. v. 1. UFMG. Belo Horizonte. p. 31-41.

PIERREHUMBERT, Janet B. (2001). Exemplar dynamics: word frequency, lenition and contrast. *In*: Bybee & P. Hopper (eds). *Frequency and the emergency of linguistic structure*. Amsterdam: John Benjamins, 137-157.

_____ (2003). Probabilistic phonology: discrimination and robustness. *In:* BOD, Rens; HAY, Jennifer; JANNEDY, Stefanie (ed.). *Probabilistic linguistics*. Cambridge, Massachusets: The MIT Press.

VIEGAS, M. C. (1987). *Alçamento das vogais pretônicas*: uma abordagem sociolinguística. MA Dissertation. UFMG. Belo Horizonte.

Prosodic Variation within Speech Planning and Execution

Vikram Ramanarayanan.
vramanar@usc.edu

Understanding prosodic variation is critical to an overall understanding of speech planning and execution and in turn, better designs of computer speech recognizers and synthesizers. Work presented here aims to examine two main facets of prosodic variation in speech production – pausing behavior and speaking style – using both the acoustic speech signal as well as data on articulation obtained using modalities such as audio-synchronized real-time magnetic resonance imaging (rt-MRI) [1,2]. The conceptual framework of Articulatory Phonology [3] and Task Dynamics [4] in conjunction with the prosodic-gesture model [5] of phrase boundaries is used as the theoretical backbone of these studies.

Speech styles are one of the primary phenomena of prosodic variation in speech. A novel automatic procedure to analyze real-time magnetic resonance images (rt-MRI) of the human vocal tract recorded for read and spontaneously spoken speech has been applied to rt-MRI data on nasal articulation, jointly used with acoustic analyses of the speech signal, to analyze nasal production differences in read and spontaneous speech, especially focusing on reduction [6]. In this exploratory study, vowel-nasal-vowel (VNV) sequences from one speaker were examined and measures extracted from both acoustic and articulatory signals. Significant differences were observed in the realizations of constriction-forming events for read and spontaneous speaking styles, and possible non-linearly related (or unrelated) characteristics of reduction were identified. Future work in this regard includes extension of this kind of reduction study to sounds produced with different manners of articulation.

Pausing in natural speech is yet another rich source of information regarding prosodic variation during spontaneous speech production. This can be considered from a listener perspective —how do pauses aid or impair speech understanding—or from a speaker perspective—how do pauses reflect the speech planning process, either operating well or encountering difficulties. It is hypothesized that pauses at major syntactic boundaries (i.e., grammatical pauses), but not ungrammatical (e.g., word search) pauses, are planned by a high-level cognitive mechanism that also controls the rate of articulation around these junctures [7]. RT-MRI is used to analyze articulation at and around grammatical and ungrammatical pauses in spontaneous speech. Measures

quantifying the speed of articulators were developed and applied during these pauses as well as during their immediate neighborhoods.

Grammatical pauses were found to have an appreciable drop in speed at the pause itself as compared to ungrammatical pauses. This is consistent with our hypothesis that grammatical pauses are indeed choreographed by a central cognitive planner, while ungrammatical pauses, which show only a slight decrease in articulator speed at the pause but are followed by a spurt in speed immediately after, give an indication of how much time it takes to 'recover' from the perturbation of the linguistic structural integrity of the utterance.

References:

[1] Narayanan, S., Nayak, K., Lee, S., Sethy, A., and Byrd, D. (2004). An approach to real-time magnetic resonance imaging for speech production. Journal of the Acoustical Society of America. 115(4): 1771-1776.

[2] Bresch, E., Nielsen, J., Nayak, K., and Narayanan, S.. (2006). Synchronized and noise-robust audio recordings during realtime MRI scans. Journal of the Acoustical Society of America.120(4): 1791-1794.

[3] Browman, C. P., & Goldstein, L., (1992) "Articulatory phonology: an overview," Phonetica, 49, 155-180.

[4] Saltzman, E. L., & Munhall, K. G. (1989). "A dynamical approach to gestural patterning in speech production," Ecological Psychology, 1, 333–382.

[5] Byrd, D. & Saltzman, E. (2003). "The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening," Journal of Phonetics, 31, 149-180.

[6] Ramanarayanan, V., Byrd, D., Goldstein, L. and Narayanan, S.. A joint acoustic-articulatory study of nasal spectral reduction in read versus spontaneous speaking styles, *Speech Prosody, Chicago, Illinois*, May 2010(accepted).

[7] Ramanarayanan, V., Bresch, E., Byrd, D., Goldstein, L., Narayanan, S. (2009). Analysis of pausing behavior in spontaneous speech using real-time magnetic resonance imaging of articulation. Journal of the Acoustical Society of America. 126 (5): EL160-EL165.

The production of [v] as an allophone of /b/ in pre-literate children in the province of Concepción (Chile)

Viviana Vergara Fernández.
vivivergaraf@gmail.com

The existence of the voiced labiodental fricative [v] as an allophone of the voiced bilabial stop /b/ in Spanish is broadly rejected. It is argued that [v] was lost in the early stages of the transformation of Latin into Spanish, and that currently it only appears as a result of assimilation, hyper-correction or inter-language interference in the case of bilinguals.

In the last several years, however, the existence of [v] as a natural allophone of /b/ in Chilean Spanish has recently begun to gain acceptance.

In order to confirm the existence of [v], and to determine its frequency and that of the other allophones of /b/, 9 native-speakers of Chilean Spanish between 4 and 5'3 years old were videotaped as they named objects presented to them as images. The recordings were analyzed visually in order to determine the point and mode of articulation of each phone ([b], [β], [v], etc.), and this data was analyzed in light of the phonetic environment in which each token occurred.

It was determined that of the 571 tokens of /b/ that were elicited, the voiced labiodental fricative [v] accounted for 49% of all tokens, followed by the voiced bilabial fricative [β] (26%), the voiced bilabial stop [b] (18%), and other allophones (7%), a category which includes the voiceless bilabial stop [p], the voiced labialized velar approximant [w], the voiced dental stop [d], the voiced dental fricative [d], the voiced labiodental approximant [ʋ], the voiced velar stop [g], the voiced velar fricative [ɣ], the voiceless bilabial fricative [ɸ], and two co-articulations, [p͡v] and [b͡v], as well as some cases of elision.

While [v] is overall the most frequent allophone of /b/, it was observed that this tendency is reversed before and after the close back vowel /u/, and after the bilabial nasal [m].

The visual analysis performed confirms recent findings that the voiced labiodental fricative [v] does indeed exist in Chilean Spanish. Furthermore, the presence and extremely high frequency of [v] in a population as young as the one studied (average age: 4 years 5 months) shows that this phone is indeed a natural allophone of /b/.

In addition, the fact that the population in question is pre-literate shows that, contrary to what some authors have postulated, the acquisition of [v] is not related to or influenced by the

grapheme <v>, and does not arise, as others have hypothesized, due to the influence of school teachers with a hyper-corrective tendency to produce [v] when teaching and correcting the pronunciation of individual words.