International Workshop on

Cross-Language Speech Perception

30-31 January 2014
University of Minho
Braga - Portugal

January 30, 2014

14:00 Registration

14:30 Opening Ceremony

14:45 Plenary Session – Denise C. Kluge: *The role of visual cues in the perception of non-native sounds*

15:45 Session: Ongoing Research (1)

17:05 Coffee break

17:30 Session: L2 Pronunciation Teaching and Miscellanea

19:30 Conference Dinner (not included in the registration fee)

January 31, 2014

09:30 Plenary Session – Ocke-Schwen Bohn: *Cross-language speech perception: What we know, what we'd like to know*

10:30 Session: Non-native Speech (1)

12:00 Lunch

13:30 Session: Ongoing Research (2)

14:50 Session: Non-native Speech (2)

16:20 Coffee break

16:45 Keynote Speaker: Márcia C. Zimmer: *Some psycholinguistic considerations on methodological issues related to speech perception tests*

17:45 Discussion and meeting about possible partnerships

18:15 Closing Ceremony

PROGRAM

Thursday, January 30

AUDITORIUM (ILCH)			
Session	**Talks**		
Plenary Session	14:45 *The role of visual cues in the perception of non-native sounds* Denise Cristina Kluge		
Ongoing Research (1)	**15:45** *Crosscultural and linguistic effects on perception of European Portuguese emotional speech* Ana Nunes	**16:05** *Speaker identification in disguised voice and imitation by Brazilian Portuguese listeners* Denise Kluge, Maria Lúcia de Gomes Castro	**16:25** *Listener: a pronunciation training system for Brazilian-accented English* Gustavo A. Mendonça, Sara Candeias, Aldebaro Klautau, Sandra M. Aluísio	**16:45** *Perception and production of European Portuguese sounds by Chinese learners* Diana Oliveira
17:05-17:30 - *Coffee break*			
L2 Pronunciation Teaching and Miscellanea	**17:30** *Verbotonal method in phonetic correction of palatal lateral approximant /ʎ/* Arnalda Dobric	**18:00** *Children's auditory perceptual performance in the identification of phonic contrasts* Larissa Cristina Berti	**18:30** *Speech technology products derived from the dialog between linguists and engineers* Andréia Schurt Rauber

Friday, January 31

AUDITORIUM (ILCH)			
Session	**Talks**		
Plenary Session	9:30 *Cross-language speech perception: What we know, what we'd like to know* Ocke-Schwen Bohn		
Nonnative Speech Perception (1)	**10:30** *Language-specific processing of speech and non-speech* Anita Wagner, Paul Iverson	**11:00** *Inhibitory control and L3 phonological processing in Catalan/Spanish bilingual learners of English* Joan C. Mora Isabelle Darcy	**11:30** *Loanword adaptation is cross-language speech perception* Silke Hamann, Paul Boersma
12:00-13:30 - Lunch			
ROOM 1005 (ILCH)			
Ongoing Research (2)	**13:30** *Perception and production of L2 English consonants and vowels: A cross training study in progress* Angelica Carlet	**13:50** *Cross-language categorization of English vowels by European Portuguese perceivers* Anabela Rato	**14:10** *Assimilation of initial English consonants to Danish* Camilla Søballe Horslund, Anne Aarhøj Andersen, Ocke-Schwen Bohn	**14:30** *Moving away from repetitions and discriminations* Yolanda Joy Calvo Benzies

Non-native Speech Perception (2)	**14:50** *The role of language proficiency in the perception of L2 voicing contrasts*	**15:20** *The perception of Western Andalusian Spanish aspirated stops by General American English listeners*	**15:50** *Perceptual assimilation of English vowels and diphthongs by Chilean Spanish speakers*
	Grzegorz Aperliński, Aleksandra Leończyk	Maria Del Saz	Juli Cebrian
	16:20-16:45 *Coffee break*		
Plenary Session	**16:45** *Some psycholinguistic considerations on methodological issues related to speech perception tests* Márcia Cristina Zimmer		
	17:45-18:15 *Discussion and meeting about possible partnerships*		

ABSTRACTS

Thursday, January 30

KEYNOTE SPEAKER
The role of visual cues in the perception of non-native sounds

Denise Cristina Kluge (Federal University of Paraná, Brazil)
deniseckluge@gmail.com

Considering that human speech is a multimodal function and can be apprehended by visual and auditory means (Rosenblum, 2005), this talk will present and discuss the role of visual cues in speech perception considering some of the main studies in the area (e.g., McGurk & McDonald 1976; Grant & Seitz, 1998). This presentation will also review and discuss some studies that investigate the effect of visual cues in the identification and/or discrimination of nonnative contrasts (Hazan, Sennema, Iba & Faulkner, 2005; Hazan, Sennema, Faulkner, Ortega-Llebaria, Iba & Chung 2006; Kluge, Reis, Nobre-Oliveira & Bettoni-Techio, 2009; Kluge, 2009), and will also address some methodological issues concerning perception tests that make use of visual input.

SESSION: ONGOING RESEARCH (1)

Crosscultural and linguistic effects on perception of European Portuguese emotional speech

Ana Nunes (University of Macau, China)
ananunes@umac.mo

Research has shown that emotions are not equally recognized. Perceptual tests demonstrated, for example, identification problems between the pairs neutral/sadness and joy/anger. Joy, sadness and anger can be considered basic emotions, as they are easier to identify perceptually across different languages and cultures. Other emotions can be part of the specificities of a Language or of an individual speaker. In order to have a more detailed picture of cross-linguistic effects on emotional speech perception, studies of additional languages are needed. As a contribution to the area, a perception test was performed using European Portuguese emotional speech. The influence of language and previous knowledge of cultural aspects was investigated. A new corpus for European Portuguese was recorded having an actor as informant. In each of two sessions, five repetitions of two semantically emotion neutral sentences (a long and a short sentence) were recorded using five emotions: joy, sadness, despair, fear, anger, and neutral. Three different groups participated in the perception (identification) test: two Chinese groups (one formed by Mandarin native speakers and the other by Cantonese speakers), both were learning Portuguese for three years; a Portuguese group, all Portuguese native speakers. Scores were compared across language, groups and emotions. Effects of linguistic factors, such as sentence length, were also analyzed. Main results show that native Portuguese speakers were better at identifying all emotions, especially sadness. All, except fear, showed statistically significant differences between the groups; for all the groups anger is the most identifiable emotion; longer sentences increase identification scores in the three groups.

Audibert, Nicolas; Aubergé, Véronique; Rilliard, Albert, "The Prosodic Dimensions of Emotion in Speech: the Relative weights of Parameters", *InterSpeech*, 2005.

Campbell, Nick; Mokhtari Parham, *Voice Quality: 4th Prosodic Dimension*, 15th ICPhS, 2003.

Cahn, Janet E., *Generating Expression in Synthesized Speech*. Master's Thesis, Massachusetts Institute of Technology, 1989.

Cowie, E. Douglas; Cowie, R. and Schroeder, M., "The description of naturally occurring emotional speech", *15th ICPhS*, 2003.

Johnstone, T. & Scherer, K. R. (1999). "The effects of emotions on voice quality". *Proceedings of the XIVth International Congress of Phonetic Sciences.*

Nunes, Ana Margarida Belém, Voice Quality in European Portuguese, PhD Thesis, University of Aveiro, Portugal, 2009.

Nunes, Ana Margarida Belém, Coimbra, Rosa Lídia, Teixeira, António, Voice Quality of European Portuguese Emotional Speech, Artificial Intelligence (LNAI) volume, 6001, 2010, 145-151, Springer 2010.

Sawamura, Kanae; Dang, Jianwu; Akagi, Masato; Erickson, Donna; Li, Aijun; Sakuraba, Kyoko; Nimematsu, Nobuaki; Hirose, Keikichi, *Common Factors in Emotion Perception Among Different Cultures,* ICPHS XVI, 2007.

Susca, Michael & Healey, E. Charles (2002). "Listener perceptions along a fluency–disfluency continuum: A phenomenological analysis". *Journal of Fluency Disorders*, 27, 135-161.

Rodrigues, Américo, "As emoções na fala [The Emotions in Speech]", Tese de Mestrado, Universidade de Aveiro, 2007.

Sawamura, Kanae; Dang, Jianwu; Akagi, Masato; Erickson, Donna; Li, Aijun; Sakuraba, Kyoko; Nimematsu, Nobuaki; Hirose, Keikichi, *Common Factors in Emotion Perception Among Different Cultures,* ICPHS XVI, 2007.

Speaker identification in disguised voice and imitation by Brazilian Portuguese listeners

Denise Cristina Kluge (Federal University of Paraná, Brazil)
Maria Lúcia de Gomes Castro (Technical Federal University of Paraná, Brazil)
deniseckluge@gmail.com, malucastrogomes@terra.com.br

This study will present preliminary results in the perception of disguised voice and imitation and it is part of a research project that aims at describing four acoustic correlates of vowel identity in two different voice qualities: normal and disguised. The present study aims at investigating whether a group of ten Brazilian Portuguese listeners can identify/discriminate speakers in normal and disguised voice, as well as in imitation. In order to do so, four discrimination tests and one identification test were designed using the free software TP (Rauber, Rato, Kluge, Santos, 2012). The disguised voice's stimuli were recorded by four men and four women reading a small text that simulated a telephone call in a kidnapping situation. Speakers were asked to read the text first two times at their normal voice, and then two times at a freely chosen form of disguise. The stimuli for the tests regarding imitation were recorded by a professional imitator. Preliminary results from a statistical analysis showed that it was more difficult for the ten listeners to identify than to discriminate the original voice from the imitation. As regards disguised voice, the results showed that it was more difficult for the listeners to discriminate the speaker when lip protrusion was the strategy used to disguise the voice.

Rauber, A. S.; Rato, A.; Kluge, D. C., Santos, G. R. dos (2012). TP – Testes de Percepção/Tarefas de Treinamento Perceptual [Perception Tests/Perceptual Training Tasks], computer program. Available at http://www.worken.com.br/tp_regfree.php

Listener: a pronunciation training system for Brazilian-accented English

Gustavo A. Mendonça[1], Sara Candeias[2], Aldebaro Klautau[3], Sandra M. Aluísio[1]
[1]University of São Paulo (ICMC/NILC/USP, Brazil)
[2]University of Coimbra (IT/UC, Portugal)
[3]Federal University of Pará (LaPS/UFPA, Brazil)
gustavoauma@gmail.com, saracandeias@co.it.pt, a.klautau@ieee.org, sandra@icmc.usp.br

Recent surveys have shown that Brazil is among the countries with the lowest knowledge of the English language. In Education First's *English Proficiency Index* (2012), for instance, Brazil ranked 46th out of 54 countries. In GlobalEnglish (2013), Brazil achieved a similar position, being ranked as 71st of 77 positions. We aim at improving the proficiency on the English language among Brazilians, by building an automatic-speech-recognition-based Pronunciation Training System for Brazilian-accented English. This Pronunciation System, called *Listener*, will be able to provide online feedback regarding the pronunciation of the user. Nine types of mispronunciations were selected to be dealt by the system, according to the study by Zimmer (2004) and those reported in Zimmer *et al.* (2009), which describe the pronunciation of English as a foreign language by undergraduate Brazilian students, classified in four levels of proficiency. We intend to present the *Listener's* prototype. It was implemented and tested, using the *Hidden Markov Model Toolkit* (HTK) (Young & Young, 1994), and applies Hidden Markov Models (HMM) to represent the acoustic features of speech. The HMMs are continuous-density Gaussian mixture tied-state triphones with clustering performed using phonetic decision trees. An excerpt (~2 hours) of the *Corpus Oral de Brasileiros Aprendizes de Inglês* (COBAI) was used to build the acoustic model. The system was tested in an isolated words task, by using forced alignment via Viterbi's algorithm. Pronunciation variants were added to the pronunciation model through symbolic rules. Some preliminary results related to consonant cluster simplification ("study" > [is.tʌdɪ]) and schwa paragoge ("book" > [bʊ.kɪ]) are analyzed and discussed. Results show that the ongoing method is promising and can allow a wide-ranging Computer-Assisted Pronunciation Training (CAPT) System.

EducationFirst. (2012). Retrieved Sept. 2nd, 2013, from English Proficiency Index 2012: http://www.ef.com.br/epi/downloads
GlobalEnglish (2012) The Business English Index 2012 Report: Analyzing the Trends of Global Readiness for Effective 21st-Century Communication. Brisbane: GlobalEnglish.
Young, S., Young, Sj (1994). The HTK Hidden Markov Model Toolkit: Design and Philosophy. In Entropic Cambridge Research Laboratory Ltd, v. 2, 2-44.
Zimmer, M. (2004) A Transferência do Conhecimento Fonético-Fonológico do Português Brasileiro (L1) para o Inglês (L2) na Recodificação Leitora: Uma Abordagem Conexionista. Dissertação de Doutorado. Porto Alegre: Pontifícia Universidade Católica do Rio Grande do Sul (Portuguese).
Zimmer, M., Silveira, R., & Alves, U. (2009). Pronunciation Instruction for Brazilians: Bringing Theory and Practice Together. Newcastle: Cambridge Scholars.

Perception and production of European Portuguese sounds by Chinese learners

Diana Oliveira
oliveira.diana27@gmail.com

Accurate pronunciation is a central research topic in the field of second language (L2) speech acquisition. It is widely accepted that foreign-accented production [caused by (supra)segmental deviations] derives mainly from perceptual inaccuracy rather than from articulatory difficulties [1; 2]. From childhood to adulthood, exposure to first language (L1) speech sounds reduces perceptual sensitivity to phonetic/acoustic categories with no functional status in the L1. This attunement to L1 sounds leads to a biased perception which, in turn, origins an accented production [3]. Therefore, research has focused on the perceptual phenomena underlying acquisition of L2 speech. Several studies, targeting different L1-L2 pairings, have investigated the efficiency of laboratory perceptual training. It has been shown that: i) short-term training is effective in improving listeners' perception skills [4; 5]; ii) perceptual learning is transferrable to new tokens and new talkers and long-term retention is possible [5; 6]; iii) enhancement in perception tends to correlate positively with performance in production [6; 7]. However, the success of perceptual training is known to vary according to the specific L1-L2 pair under study, depending on the phonological/phonetic and phonotactic (dis)similarities between the two languages [3]. This study aims at testing the efficiency of perceptual training in the improvement of perception and production of European Portuguese (EP) consonant sounds by Mandarin speakers, who are currently an important group of EP learners [8; 9]. To our knowledge, no experimental investigation to date has studied the combination of Mandarin as L1 and EP as L2. Thus, it is important to understand: i) which EP sounds are incorrectly produced by Mandarin speakers; ii) whether or not and to what extent previous results of successful training procedures can be accomplished for this specific L1-L2 pairing; iii) how resulting knowledge can be made useful for designing adequate EPL2 teaching materials.

1. Flege, J. (1995). Second Language Speech Learning: Theory, Findings and Problems. In Strange, W. (Ed), Speech Perception and Linguistic Experience: Issues in Cross Language Research (pp. 233-277). Timonium, MD: New York Press.
2. Best, C. & Tyler, M. (2007). Nonnative and Second Language Speech Perception: Commonalities and Complementarities. In Bohn, O. & Munro, M. (Eds), Language Experience in Second Language Speech Learning – In Honor of James Emil Flege (pp. 13-34). Amsterdam/Philadelphia: John Benjamins Publishing Company.
3. Strange, W. (1995). Cross language studies of speech perception - a historical review. In Strange, W. (Ed), Speech Perception and Linguistic Experience: Issues in Cross Language Research (pp. 3-45). Timonium, MD: New York Press.
4. Rochet, B. (1995). Perception and Production of Second Language Speech Sounds by Adults. In Strange, W. (Ed), Speech Perception and Linguistic Experience: Issues in Cross Language Research (pp. 379-410). Timonium, MD: New York Press.
5. Wang, X. (2008). Perceptual Training for Learning English Vowels – Perception, Production and Long-Term Retention. Saarbrücken, Germany: Verlag Dr. Müller.
6. Yamada, R., Tohkura, Y., Bradlow, A. & Pisoni, D. (1996). Does Training in Speech Perception Modify Speech Production?. Proceedings of the Fourth International Conference on Spoken Language Processing, v. 2 (pp. 606-609). Philadelphia: Institute of Electrical and Electronics Engineers.
7. Alliaga-García, C. & Mora, J. (2009). Assessing the effects of phonetic training on L2 sound perception and production. In Watkins, M., Rauber, A. & Baptista, B. (Eds), Recent Research in Second Language Phonetics/Phonology – Perception and Production (pp. 2-31). Newcastle upon Tyne: Cambridge Scholars Publishing.
8. Reto, L. (Coord) (2012). Potencial Económico da Língua Portuguesa. Alfragide: Texto Editores.
9. Belchior, M. (2011). Chineses aprendem mais português. Revista Macau (23). Retrieved May 24, 2013, from http://www.revistamacau.com/2011/06/05/chineses-aprendem-mais-portugues/

SESSION: L2 PRONUNCIATION TEACHING AND MISCELLANEA

Verbotonal method in phonetic correction of palatal lateral approximant /ʎ/

Arnalda Dobric (Faculty of Humanities and Social Sciences, University of Zagreb)
arnalda.dobric@gmail.com

Although often neglected in foreign language teaching, pronunciation holds a very important role in language acquisition or learning. On the phonological level adult learners of foreign languages tend to have problems with the speech sounds that are close to some sounds in their native language but learners do not recognize them as different and replace them with the similar speech sounds from their native language. These errors are almost systematic if learners belong to the same L1 and learn the same L2. In the verbotonal method it is referred to as the system of errors. Another set of problems arises with the sounds that do not exist in the learners' native language. In this paper, effects of phonetic training based on several elements (sound context with respect to optimal frequency band and coarticulation, position of

the target sound within the word or sentence, intonation, tenseness of articulation as well as movement and body posture accompanying speech production) according to the verbotonal method are shown in the example of the palatal lateral approximant /ʎ/, which is one of the hardest sounds to pronounce for foreign learners of Croatian as L2. Learners of different native languages (Spanish, Czech, English, German and Hungarian) were recorded before and after sessions of phonetic training. The results were assessed using a perceptual test. Furthermore, acoustical analyses were performed in Praat. The results showed higher F2 after correction sessions. Since body posture accompanying speech production seems to be one of the crucial elements for correction of this specific error, the importance of body movement and posture in speech is discussed in more detail. Overall changes in pronunciation after phonetic training are also discussed.

Children's auditory perceptual performance in the identification of phonic contrasts

Larissa Cristina Berti (Paulista State University, Brazil)
berti.larissa@gmail.com

The aim of this study was to assess the auditory perceptual performance of children in a task of identification of phonic contrasts in Brazilian Portuguese (henceforward BP), and to classify which phonic classes provide higher or lower degrees of difficulty. A forced-choice minimal-pair identification task involving the 19 phonemes in BP was conducted with 66 children, 4-5 years old, using PERCEVAL software. The stimuli used in the identification task consisted of a typical adult's recordings of the familiar disyllabic words (minimum pairs) contrasting the 19 phonemes in stressed position. The acoustic stimulus was presented to the children using headphones and they needed to choose which stimulus-correspondent picture was shown on the computer screen. Both presentation time and reaction time of the stimulus were measured by PERCEVAL software. Our results showed that the perceptual accuracy was dependent on phonic classes according to the following decreasing order: vowels (d' prime = 1,73) > sonorants (d' prime = 1,56) > occlusives (d' prime = 1,43) > fricatives (d' prime = 0,93). The reaction time of the incorrect response was significantly higher than the reaction time of the correct response in all the phonic classes. Within each class, there are contrasts that present a greater difficulty in the identification task. Altogether, the perceptual difficulty varies due to phonetic distance. The auditory perceptual mastering of the phonic contrasts in BP occurs gradually and, within each class, there seems to be a hierarchy in the perceptual mastering. The parallel between perceptual acquisition orders and production acquisition orders was not always the same.

Speech technology products derived from the dialog between linguists and engineers

Andréia Schurt Rauber (Appen Butler Hill)
asrauber@gmail.com

The partnership between linguists, engineers and programmers has become crucial in order to design speech products that have increasingly become part of our daily lives. Voice commands have become easily available in cars, smartphones and smart TVs, for example, but how does the machine recognize human speech, process requests and react to users? In this talk I will present linguistic tasks that are essential for the creation of speech recognizers and synthesizers, and will also stress that the knowledge of first and foreign language phonology, and (acoustic) phonetics is necessary for engineers and programmers to develop the voice-operated apps that surround us. Three questions motivate this talk: (1) are phonologists/phoneticians aware of their importance for the creation of speech products? (2) Are linguists open to interact with other fields and also go beyond academic projects? (3) What is the role of speech perception in the design of synthesizers that use the phonological inventory of a given language to synthesize words in other languages?

KEYNOTE SPEAKER
Cross-language speech perception: What we know, what we'd like to know

Ocke-Schwen Bohn (Aarhus University, Denmark)
engosb@hum.au.dk

The perception of nonnative speech sounds is guided by two kinds of biases: Those which can be attributed to the native sound system, and those which reveal universal preferences for particular acoustic dimensions or types of sounds. This presentation will briefly review the success with which the most widely used models of second language (L2) speech, Flege's SLM (1995) and Best's PAM (1995) and PAM-L2 (Best & Tyler 2007) predict the application of L1 biases to L2 perception, and it will then attempt to address open questions regarding L1 biases, such as how to assess cross-language similarity of speech sound categories, and questions surrounding the role of higher-order phonological biases in cross-language speech perception. The presentation will also review two universal biases in cross-language speech perception which, respectively, provide the basis for Bohn's (1995) desensitization hypothesis and for Polka and Bohn's (2003, 2011) Natural Reference Vowel framework. These well-documented biases inspire questions regarding less-well studied phenomena in consonant and vowel perception which might reveal universal preferences.

Best, Catherine T. 1995. A direct realist view of cross-language speech perception. In: Strange, W., ed., Speech perception and linguistic experience: Issues in cross-language research. Timonium, MD: York Press. 171-204.

Best, Catherine T. & Tyler, Michael D. 2007. Nonnative and second language speech perception. In: Bohn, Ocke-Schwen & Munro, Murray J., eds. 2007. Language experience in second language speech learning: In honor of James Emil Flege. Amsterdam & Philadelphia: John Benjamins, 13-34.

Bohn, Ocke-Schwen. 1995. Cross-language speech perception in adults: First language transfer doesn't tell it all. In: Strange, W., ed., Speech perception and linguistic experience: Issues in cross-language research. Timonium, MD: York Press, 279-304.

Flege, J. E. 1995. Second language speech learning: Theory, findings, and problems. In: Strange, W., ed., Speech perception and linguistic experience: Issues in cross-language research. Timonium, MD: York Press. 233-277

Polka, Linda & Bohn, Ocke-Schwen. 2003. Asymmetries in vowel perception. Speech Communication 41, 221-231.

Polka, Linda & Bohn, Ocke-Schwen. 2011. Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. Journal of Phonetics 39, 467-478.

SESSION: NONNATIVE SPEECH PERCEPTION (1)

Language-specific processing of speech and non-speech

Anita Wagner (University Medical Center Groningen, The Netherlands)
Paul Iverson (University College London, UK)
a.wagner@umcg.nl, p.iverson@ucl.ac.uk

Adult listeners perceive speech sounds in language-specific ways. It is an open question what levels of processing are adapted to the mother tongue, but it is generally assumed that language experience does not alter auditory processing. We investigated language-specific sensitivities to acoustic differences between English /r/ and /l/ in Japanese listeners, and studied the processing of these sounds at preattentive auditory stages and at levels of phonetic categorization. When listening to speech, Japanese listeners show language-specific sensitivities to the acoustic dimensions that differentiate English /r/ from /l/. When presented with the same acoustic variation in non-speech contexts, however, Japanese listeners show a pattern of discrimination not different from English. The present study used stimuli that parametrically varied this speech/nonspeech difference, to see how far a stimulus must be from speech in order for cross-language differences to disappear. The experimental stimuli were a series of continua that mimicked the acoustic differences between /r/ and /l/. The different continua disrupted, to varying degrees, the naturalness and perceptual coherence of the stimuli. Behavioral results demonstrated strong cross-language differences for the stimuli that sounded most like speech, no significant cross-language differences for the stimuli that were least like speech, and moderate cross-language differences for a middle range of stimuli that were not categorized accurately as /r/ and /l/ and did not sound like speech. The cross-language differences for the latter stimuli were not caused by phonological categorization, but they may be somewhat speech specific. An MEG experiment recorded listeners' mismatch responses (MMF) to four stimulus series, which showed different degrees of language-specific processing. These results show cross-language differences in the processing of all four transforms. Stimuli that defied phonetic categorization showed language-

dependent processing at a low pre-attentive level. This suggests that language-specialization affects the processing of acoustic variation for both speech and non-speech.

Inhibitory control and L3 phonological processing in Catalan/Spanish bilingual learners of English

Joan C. Mora (Universitat de Barcelona, Spain)
Isabelle Darcy (Indiana University, USA)
mora@ub.edu, idarcy@indiana.edu

Recent research suggests that L2 learners with better inhibitory control may be more successful at keeping their L1 and L2 phonological systems apart, thus avoiding cross-language phonological interference in speech perception and production (Darcy, Mora, Daidone, in press; Lev-Ari, Peperkamp, 2013). The present study investigates whether Catalan/Spanish bilingual learners of English with better inhibitory skill develop more accurate representations for L3-English sounds. Inhibition was assessed through a retrieval-induced inhibition task (Lev-Ari, Peperkamp, 2013; Veling, van Knippenberg, 2004). L3-English phonology was assessed through a speeded ABX categorization task (perception) and a delayed sentence repetition task (production), while partialling out proficiency through an L3 vocabulary size measure. Participants were early Catalan/Spanish bilinguals, who were also late learners of English (their L3). The participants used both Catalan and Spanish daily, but crucially differed in the amount of use of their less dominant language (either Catalan or Spanish). They were assigned to two groups according to the percentage of use of their less dominant language: "unbalanced" (<30%) vs. "balanced" (>30%). We predicted that the unbalanced group, unlike the balanced group, would have developed a stronger inhibitory control due to their extensive experience in inhibiting their more dominant language when using their less dominant one. We further hypothesized that their enhanced inhibitory control would be related to more efficient L3 (English) phonological processing. Results showed that the "unbalanced" group indeed had a significantly stronger inhibitory control than the "balanced" group, in the way we had predicted. However, inhibitory control scores were unrelated to L3 phonological scores. We tentatively concluded that the effects of individual differences in inhibitory control in these early bilinguals might have been "washed out" due to the daily practice they receive in inhibiting one language over the other, which calls for further research on the relationship between inhibitory control and L2 and L3 phonological processing.

Darcy, I., Mora, J. C. & Daidone, D. (*in press*). Attention control and inhibition influence phonological development in a second language. In *Proceedings of the International Symposium on the Acquisition of Second Language Speech. Concordia Working Papers in Applied Linguistics, 5.*

Lev-Ari, S., & Peperkamp, S. (2013). Low inhibitory skill leads to non-native perception and production in bilinguals' native language. Journal of Phonetics,41, 320-331.

Veling, H. & van Knippenberg, A. (2004). Remembering can cause inhibition: Retrieval-induced inhibition as cue independent process. *Journal of Experimental Psychology: Learning, Memory and Cognition* 30*, 2*, 315-318).

Loanword adaptation is cross-language speech perception: case studies from Malayalam, Thai and Japanese

Silke Hamann, Paul Boersma (University of Amsterdam, The Netherlands)
silke.hamann@uva.nl, paul.boersma@uva.nl

Studies on loanword adaptation (such as Davidson & Noyer, 1996; Broselow, 2003; Yip, 2006) usually employ loanword-specific devices to deal with what Kenstowicz (2005) called "divergent repairs": loanword adaptations differ from what we would expect on the basis of the native phonology. In the present account of divergent repairs we follow Boersma and Hamann (2009) by assuming that loanwords are integrated into a language via native-language perception of nonnative sounds and forms. This perception process involves both language-specific perceptual cues and structural restrictions, and can be formalized in a perception grammar (Boersma, 2000, see lower left of the figure) as the interaction of cue and structural constraints:
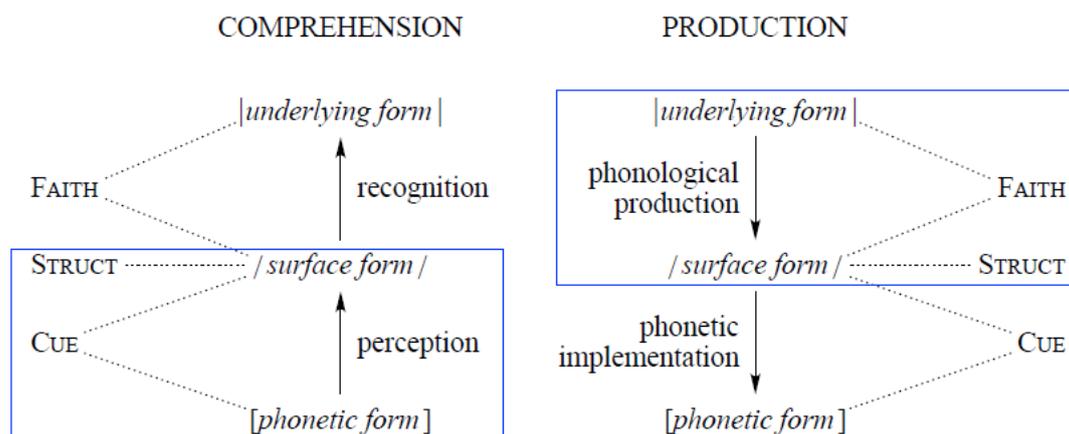
In native language phonological production (upper right part of the figure), on the other hand, the same structural constraints interact with phonological faithfulness constraints. In Malayalam (Mohanan & Mohanan, 2003), for example, the structural constraint "no voiceless singleton intervocalically", which is unviolated in the language, is satisfied by gemination in loanword perception, but by voicing in native phonological production. The English word *baker* enters the Malayalam ear as the foreign phonetic form [beikəɹ] and is converted natively to the surface structure /.beːk.kar./ because the cue constraint *[no periodicity]/voi/ ("phonetic non-periodicity does not correspond to phonological voicing") outranks the cue constraint *[small duration]/long/ ("phonetic shortness does not correspond to a phonological long segment"). In Malayalam phonological production, the same structural constraint interacts with the faithfulness constraints IDENT(length) and the lower-ranked DEP(voi), which are responsible for the realization of underlying intervocalic singleton |makan| as voiced /.ma.gan./. The divergent repairs are thus caused by the different types of constraints that the structural constraint fights with: faithfulness in production and cues in perception. We will present further illustrations of divergent repairs from Thai and Japanese, and conclude that loanword adaptation is often just cross-language speech perception.

Friday, January 31 - afternoon

SESSION: ONGOING RESEARCH (2)

Perception and production of L2 English consonants and vowels: A cross training study in progress

Angelica Carlet (Universidad Autonoma de Barcelona, Spain)
angelica.carlet@uab.cat

Foreign language learning in a formal setting is notably characterized by limited target language exposure. Thus, phonetic training may play a very important role in pronunciation learning, as it can make up for the absence of regular L2 experience. Several studies have shown the efficacy of phonetic training on the ability to correctly perceive and produce L2 consonant sounds (Hardison, 2003; Iverson & Evans, 2007; Pisoni, Lively & Logan, 1994; Strange & Dittmann, 1984, among others) and L2 vowel sounds (Aliaga-García, 2011; Aliaga-García & Mora, 2009; Ceñoz & Garcia Lecumberri, 1999; Iverson & Evans, 2007, among others). This study will further investigate the effects of a high variability phonetic training approach (Logan, Lively & Pisoni, 1991) on the perception and production of non-native sounds, specifically of the English stop consonant sounds and the English vowel sounds [/æ/-/ʌ/,/ɪ/-/i/] by Spanish/Catalan speakers. Two experimental groups (N=30) and a control group (N=20) will be exposed to a perceptual cross-training methodology and will be tested on both their perception and production of L2 sounds by means of a pretest/posttest/retention-test basis. Each experimental group will be explicitly trained (instruction and feedback) to better identify either vowels or consonants embedded in nonce words, while being implicitly exposed to the untrained segments (vowels or consonants). This study also aims at investigating if improvement generalizes to aspects not explored in the training, such as new words, new talkers and to "implicitly trained" segments. The predictions are that phonetic training will improve L2 performance for the explicitly trained segments and to a lesser extent to the "cross-trained" segments, due to the lack of focus on specific cues and feedback. Moreover, generalization is predicted to occur, as well as retention of the knowledge acquired during training.

Cross-language categorization of English vowels by European Portuguese perceivers

Anabela Rato (Universidade do Minho, Portugal)
asrato@ilch.uminho.pt

Cross-language research on phonological learning has revealed that the perception of nonnative vowel contrasts is highly challenging for adult L2/FL learners mainly due to perceptual biases caused by L1 interference [1]. Studies on English vowel categorization by perceivers from small-vowel-set romance languages such as Spanish [2, 3, 4], Italian [5, 6], and Portuguese [7, 8, 9] have shown that perceptual patterns in the mapping of English vowels onto L1 phonological categories are, to a great extent, similar. However, it is important to further investigate each language variety so that specific perceptual difficulties can be predicted accurately. Therefore, although there are some studies that investigated the categorization of English vowels by native Brazilian Portuguese (BP) listeners, research regarding the assimilation of L2 vowels by European Portuguese (EP) perceivers is scarce [10, 11]. In order to assess cross-linguistic perceptual (dis)similarity between the English and the EP vowel inventories and to further understand reported inaccuracy in the assimilation of nonnative vowel contrasts, two experiments were carried out with 68 Portuguese learners of English as a Foreign Language (EFL).The categorization of a set of English vowels was tested with two identification tasks. In Experiment 1, a group of 34 participants was tested on native and nonnative vowels with an 8AFC cross-language perceptual assimilation task (PAT), in which listeners identified the target segments in terms of L1 vowel categories and then rated each segment for goodness-of-fit. In Experiment 2, the other group of 34 students was tested exclusively on the set of nonnative vowels with a 7AFC ID task, by labelling them according to the target L2 categories. In this paper, the results of the two experiments are discussed by underlying the L1-L2 (dis)similarities, and a comparison amongst the results of the aforementioned studies and the findings of these two experiments is made.

1. Strange, W. (2007). Cross-language phonetic similarity of vowels: Theoretical and methodological issues. In O.-S. Bohn, & M. Munro (Eds.), *Language Experience in Second Language Speech Learning – In honor of James Emil Flege* (pp. 35-55). Amsterdam/Philadelphia: John Benjamins Publishing Company.
2. Cebrian, J. (2006). Experience and the Use of Duration in the Categorization of L2 Vowels. *Journal of Phonetics*, *34*, 372-387.
3. Flege, J., Bohn, O.-S. & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, *25*, 437-470.
4. Flege, J., Munro, M., & Fox, A. (1994). Auditory and categorical effects on cross-language vowel perception. *Journal of the Acoustical Society of America*, *95*(6), 3623-3641.
5. Flege, J., MacKay, I., & Meador, D. (1999). Italian speakers' perception and production of English vowels. *Journal of the Acoustical Society of America*, *106*(5), 2973-4.
6. Flege, J., & MacKay, I. (2004) Perceiving vowels in a second language. *Studies in Second Language Acquisition*, *26*, 1-34.
7. Major, R. (1987). Phonological Similarity, Markedness, and Rate of L2 Acquisition. *Studies in Second Language Acquisition*, *9*(1), 63-82.
8. Nobre-Oliveira, D. (2007). *The Effect of Perceptual Training on the Learning of English Vowels by Brazilian Portuguese Speakers* (Doctoral dissertation).
9. Rauber, A. S. (2010). *Acoustic characteristics of Brazilian English vowels: perception and production results*. Saarbrücken: Lambert Academic Publishing.
10. Flege, J. (1995). Second Language Speech Learning: Theory, Findings and Problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross Language Research* (pp. 233-277). Timonium, MD: York Press.
11. Rato, A., Rauber, A., Soares, L., & Lucas, L. (2013). *Challenges in the perception and production of English front vowels by native speakers of European Portuguese*. Manuscript submitted for publication.

Assimilation of initial English consonants to Danish: Identifying areas of expected learning ease and difficulty for native Danish speaking learners of English

Camilla Søballe Horslund, Anne Aarhøj Andersen,
Ocke-Schwen Bohn (Aarhus University, Denmark)
engcsh@dac.au.dk, stud20115744@hum.au.dk, engosb@hum.au.dk

Several studies suggest that second language (L2) learners face different challenges depending on how their native language (L1) maps onto the target language (e.g., Flege et al., 1997), and hence leading models of L2 speech perception and production take L1-L2 assimilation as their starting point. The Speech Learning Model (SLM) predicts learning success for the production of individual L2 phones based on their perceived similarity to phones in the learner's L1, dividing the L2 phones into new, similar and identical phones vis-á-vis the L1 (Flege 1995). Likewise, the Perceptual Assimilation Model for L2 learners (PAM-L2) predicts discrimination performance of pairs of L2 sounds based on how the relationship between the two L2 phones is assimilated to the native phoneme inventory; i.e., whether the two L2 phones are assimilated to the same L1 category or to different L1 categories as well as the L2 phones' perceived goodness of fit as instances of the L1 category (Best & Tyler, 2007). This study investigates how initial English consonants are

assimilated to the phoneme inventory of Danish. Twenty initial English consonants were presented in a /Ca/ frame and listeners were asked to map them onto one of 15 response alternatives representing initial Danish consonants and consonant clusters. Listeners also rated the goodness of fit of each assimilation on a 9-point Likert scale. The results of the study thus present the first step towards predicting learning ease and difficulty of individual English consonants for L1 Danish learners of English. The talk will present the assimilation results as well as categorise initial English consonants in terms of SLM and PAM-L2 for L1 Danish learners of English. The talk will also present hypotheses regarding the acquisition of initial English consonants by L1 Danish speakers based on these categorisations.

Flege, J. E. 1995. Second language speech learning: Theory, findings, and problems. In: Strange, W., ed., *Speech perception and linguistic experience: Issues in cross-language research.* Timonium, MD: York Press. 233-277.

Flege, J.E., Bohn, O.-S. and Jang, S. 1997. Effects of Experience on Non-Native Speakers' Production and Perception of English Vowels". *Journal of Phonetics.* 25, 437-470.

Best, C. T. & Tyler, M. D. 2007. Nonative and second language speech perception. In: Bohn, O.-S. & Munro, M. J., eds. 2007. *Language experience in second language speech learning: In honor of James Emil Flege.* Amsterdam: John Benjamins, 13-34.

Moving away from repetitions and discriminations: introducing engaging and motivating activities into EFL pronunciation classes to help Spanish students improve their perception and production of English vowels

Yolanda Joy Calvo Benzies (Universidad de Compostela)

yolandajoy.calvo@rai.usc.es

Spanish learners of English, even those at advanced levels, tend to have serious problems with the oral component and, more particularly, with pronunciation (Alcaraz & Moody, 1999; Palacios, 2000; Martínez, Usó, &Alcón, 2006). Generally speaking, these difficulties are due to the many differences between the phonological systems of both languages and to the low degree of exposure that students from this country have to spoken English outside the classroom. Hence, it would seem feasible for pronunciation to be emphasized in English as a foreign language (EFL) classes in Spain. However, even though students consider that pronunciation should receive more attention in their EFL classes (Calvo, 2013a, b), the truth is that this language area continues to receive less attention than other language fields such as grammar or vocabulary. Moreover, the majority of activities present in EFL textbooks addressed at Spanish learners are mainly simple drills, repetitions and discrimination exercises (Calvo, 2013c), without engaging the students into real and authentic oral communication. This paper aims at helping Spanish students improve both their perceptual and productive levels of pronunciation by: a) describing some of the most modern and authentic materials available for teaching foreign languages such as restaurant menus, podcasts, maps and songs; and, b) providing ideas and activities to generally focus on oral skills and, more specifically, on the pronunciation of English vowels, by using the materials mentioned above. These ideas will address the main difficulties that Spanish learners tend to have with English vowels, such as the distinctions between short and long vowels, schwa or /^/; authentic materials and topics that students are likely to use in daily-life situations, such as, food, health and travelling, have been chosen so as to emphasise real oral communication, rather than the artificial dialogues generally present in EFL materials addressed at Spanish learners.

Alcaraz, Enrique and Moody, Brian (1999). Fonética inglesa para españoles. Alcoy: Marfil.

Calvo Benzies, Yolanda Joy (2013a). "Not Enough Time is Devoted to Pronunciation, I would Like my Textbook to Include other Types of Pronunciation Activities. The Views and Perspectives of EFL Fourth Year Secondary-school Students on the Teaching and Learning of Pronunciation. A Survey-based Study" in Mompeán, J.A. and Fouz, Jonás (eds.), Proceedings of the Third International Conference on English Pronunciation: Issues & Practices (EPIP3). Murcia: Universidad de Murcia, pp. 10-13.

Calvo Benzies, Yolanda Joy (2013b). "Spanish EFL University Students' Views on the Teaching of Pronunciation: A Survey-Based Study" in Ciarlo, C and Giannoni, D.S (eds.). Language Studies Working Papers, vol. 5. Reading: University of Reading, pp. 41-49.

Calvo Benzies, Yolanda Joy (2013c). "La enseñanza de la pronunciación del inglés en Galicia: una asignatura pendiente" in Cabedo Nebot, A.; Aguilar Ruiz, M.J; and López-Navarro Vidal, E. (eds.), Estudios de Lingüística: investigaciones, propuestas y aplicaciones. Valencia: Universitat de Valencia, pp. 13-20.

Martínez-Flor, A., Usó-Juan, E. and Alcón Soler, E (2006). "Towards acquiring communciative competence through speaking" in Usó-Juan, E. and Martínez-Flor, A. (eds). Current trends in the development and teaching of the four language skills. Berlin: Mouton de Gruyter, pp, 139-157.

Palacios Martínez, I.M (2000). "Improving learner´s pronunciation of English: some reflections and some practical tips" in Ruiz, José Mª; Sheerin, Patrick H. and Estebánez, Cayetano (eds.). Estudios de metodología de la lengua inglesa, pp, 17-40.

The role of language proficiency in the perception of L2 voicing contrasts

Grzegorz Aperliński (Adam Mickiewicz University, Poznań)
Aleksandra Leończyk (University of Szczecin)
gaperlinski@wa.amu.edu.pl, leo.aleksandra@gmail.com

English and Polish discriminate between voiced and voiceless pairs of consonants, mostly on the basis of voice onset time (VOT). English is an aspirating language (Lisker & Abramson, 1964), whereas Polish is a voicing language (Keating, 1980). Taking this into consideration, one might conclude that English voiced and Polish voiceless obstruents should sound the same to a naïve Polish listener. A previous study on the perception of voicing revealed that when English voiceless and voiced obstruents were acoustically manipulated to have the same VOT, Polish listeners correctly identified English voiced obstruents in 80% of the cases, while English listeners did so nearly 100% of the time (Aperliński, 2012). While Polish listeners were shown to differ significantly from English listeners in terms of voicing perception, no difference was found for Polish listeners with different proficiency levels (beginner vs. advanced). The authors believe that the observed tendency might not have been necessarily due to the lack of differences between the groups but due to methodological considerations, i.e., the use of self-reported proficiency. The current study aims at further exploring the L2 perception of voicing with an improved method for testing the participants' proficiency. Over 60 participants divided into two groups differing significantly in their language proficiency performed an identical task to that in Aperliński (2012), i.e., forced-choice identification task on 21 sets of English minimal pairs, e.g., pit/bit. The tokens were acoustically manipulated so that both words in a set had the same short lag VOT. However, the careful control over their language skills (testing rather than self-reporting) should provide better insights into the role of language proficiency. If highly proficient Polish listeners display tendencies similar to those of native speakers of English, the observed results might be interpreted as evidence to the acquisition of L2 phonology.

Aperliński, G. 2012. Is VOT enough? (Paper presented at the VIIth International Conference on Native and Non-native Accents of English, 6-8 December 2012.).
Keating P. A. 1980. A Phonetic Study of a Voicing Contrast in Polish [Unpublished Ph.D. dissertation, Brown University.].
Lisker, L. and Abramson, A. S. 1964. A cross-language study of voicing initial stops: Acoustic measurements. *Word* 21: 384–422.

The perception of Western Andalusian Spanish aspirated stops by General American English listeners

Maria Del Saz (Universidad de Sevilla, Spain)
msdelsaz@gmail.com

In Western Andalusian Spanish (WAS), syllable- and word-final -*s* is realized as aspiration [h]. This phenomenon is not limited to southern Spain; it is also present in other areas such as the Canary Islands or the Caribbean and the Pacific countries. After aspiration, voiceless stops become aspirated [pʰ, tʰ, kʰ] (Alvar, 1996; Gerfen, 2002; Torreira, 2007a, 2012). In the case of WAS, these stops are characterized by longer post-aspiration and shorter pre-aspiration than in other Spanish aspirating dialects (Torreira, 2007b). This study explores the perception of aspirated –*s* before this set of stops by 84 General American English (GAE) listeners, learners of Spanish, and aims at pointing the acoustic characteristics that favor the identification of such underlying –*s*. Stimuli consisted of nouns and verbs ending in the morphological marker –*s* or in vowel, followed by different phonetic contexts, embedded in sentences. These sentences were presented randomly in an identification task with two forced-choice alternatives. Results of the identification of aspirated –*s* before the voiceless stops indicate that listeners identified [tʰ] significantly better than [kʰ] (Z = -4.696, p < .001), which in turn was better identified than [pʰ] (Z = -3.826, p < .001). An acoustic analysis of the target stops in terms of previous vowels (pre-aspiration) and Voice Onset Time (VOT) (postaspiration) revealed no significant differences in VOT duration between the three aspirated stops [χ2(2) = -.512, p = .77], while significant differences in the duration of their preceding vowels appeared [χ2(2) = 6.03, p < .05], which were longer before [tʰ] than before the other two stops. Conversely, vowels before unaspirated [t] were particularly shorter than before the other two unaspirated stops [χ2(2) = 6.53, p < .05]. From this analysis, it seems that the longer pre-aspiration observed before [tʰ] was a key factor in identification.

Alvar, M. (1996) *Manual de Dialectología Hispánica. El Español de España*. Barcelona: Ariel.
Gerfen, C. (2002) Andalusian Codas. *Probus, 14,(2),* 247-277.
Torreira, F. (2007a) Pre- and post-aspirated stops in Andalusian Spanish. In Pilar Prieto, Joan Mascaro & Maria-Josep Solé (Eds.), *Prosodic and Segmental Issues in Romance*, (pp. 67–82). Amsterdam: John Benjamins.

Torreira, F. (2007b) Coarticulation between aspirated-s and voiceless stops in Spanish: An interdialectal comparison. In N. Sagarra & A. J. Toribio (Eds.), *Selected Proceedings of the 9th Hispanic Linguistics Symposium*, (pp. 113–120). Somerville, MA: Cascadilla Press.

Torreira, F. (2012) Investigating the nature of aspirated stops in Western Andalusian Spanish. *Journal of the International Phonetic Association, 42*, 49-63.

Perceptual assimilation of English vowels and diphthongs by Chilean Spanish speakers

Juli Cebrian (Universidad Autonoma de Barcelona, Spain)
juli.cebrian@uab.es

The degree of similarity between existing native language (L1) categories and second/foreign language (L2) sounds plays a crucial role in the categorization of target L2 sounds. In some cases, L1-L2 differences in inventory size result in two or more L2 sounds mapping onto a single L1 category. Spanish and English differ greatly in vowel inventory size, and cross-language comparisons typically involve the whole set or a subset of the English vowels and the five Spanish monophthongs (e.g., Flege, Munro & Fox, 1994; Imai, Flege & Wayland, 2002; Iverson & Evans, 2007). However, such an approach ignores the possible role of Spanish diphthongs in crosslinguistic perception (cf. Escudero & Williams, 2011). Further, the choice of target language variety may influence the results of cross-language perceptual mapping (e.g., Escudero & Chládková, 2010). The goal of the current study is to evaluate the role of Spanish diphthongs in the perceptual assimilation of English vowels to Spanish vowels. The study also explored if cross-language mappings are influenced by the choice of target language variety. A group of 27 Chilean Spanish-speaking learners of English performed two perceptual assimilation tasks in which listeners identified Southern British English and Canadian English stimuli in terms of Spanish vowel categories and provided goodness-of-fit ratings. Results showed that, while a few pairs of English vowels were mapped onto the same Spanish vowel (e.g., English /i-ɪ/-Spanish /i/), each English diphthong tested was consistently assimilated to a Spanish diphthong rather than to a monophthong. Despite common general trends across the two varieties of English, some differences also emerged. Taken together, the results show that vowel sequences like Spanish diphthongs should be included in cross-linguistic comparison studies, and that cross-language similarity mappings may vary depending on the target language variety tested. Possible talker effects and additional language situations are also discussed.

Escudero, P., Chládková, K. (2010). Spanish listeners' perception of American and Southern British vowels. *J. Acoust. Soc. Am. Express Letters* 128(5), EL254-260.

Escudero, P., Williams, D. (2011). Perceptual assimilation of Dutch vowels by Peruvian Spanish listeners. *J. Acoust. Soc. Am. Express Letters* 129(1), EL1-7.

Flege, J.E., Munro, M.J., Fox, R.A. (1994). Auditory and categorical effects on cross-language vowel perception. *Journal of the Acoustical Society of America*, 95: 3623-3641.

Imai, S., Flege, J. E., Wayland, R. (2002). Perception of cross-language vowel differences: A longitudinal study of native Spanish learners of English. *Journal of the Acoustical Society of America,* 111, 2364.

Iverson, P., Evans, B.G. (2007). Learning English vowels with different first-language vowel systems: Perception of formant targets, formant movement, and duration. *Journal of the Acoustical Society of America* 122(5), 2842-2854.

KEYNOTE SPEAKER
Some psycholinguistic considerations on methodological issues related to speech perception tests

Márcia Cristina Zimmer (Ritter dos Reis University Center, Brazil)

This talks departs from the concept and phenomenon of categorical perception in neuroscience studies of speech perception to review controversial methodological issues involved in speech perception tests, such as AX, Speeded AX, ABX (and its variants AXB, XAB), 2AFC, 4IAX. Some of the tests' main psycholinguistic constraints reported in the literature are related to memory taxing and recency effects. However, some of the advantages of these tests are also highlighted.