LabPhon 9
Change in Phonology

Ninth Conference on Laboratory Phonology
Thursday 24 June - Saturday 26 June, 2004

LabPhon 9 hosts

Location

Important Dates

Program Committee

Conference Grants

Downloads

- **Schedule [pdf]**
- **Shuttle Service**
- **Style Guidelines [pdf]**

Please follow this link to the visit the 10th Conference on Laboratory Phonology page: *Variation, Detail and Representation*

Welcome to the 9th Conference on Laboratory Phonology

LabPhon is an international forum for interdisciplinary research on the sound structure of languages. Participants in LabPhon seek to communicate across the traditional boundaries that separate phonology, as a branch of theoretical linguistics, from the study of speech production, speech perception, spoken language acquisition, computer speech processing, and other disciplines concerned with human speech.

The theme for LabPhon 9, ***Change in Phonology***, addresses questions related to the evolution of language within a speech community and the development of language within the individual speaker/hearer, and includes issues of broad social interest, such as child language development, language pedagogy, and technologies for human-computer interaction.

Themes and Invited Participants

Acquisition as change: L1 phonology
LouAnn Gerken, University of Arizona , invited speaker
Stefan Frisch, University of South Florida, discussant

Phonological models of variation in computer speech processing
Julia Hirschberg, Columbia University, invited speaker
Carol Espy-Wilson, University of Maryland, discussant

Prosodic influence on change in sound patterns
Cécile Fougeron, CNRS-Université Paris 3, invited speaker
Kenneth de Jong, Indiana University, discussant

Social factors in phonetic variation
Gerry Docherty, University of Newcastle upon Tyne, invited speaker
Henrietta Jonas-Cedergren, Université du Québec à Montréal, discussant

Mechanisms of sound change
Jonathan Harrington, University of Kiel, invited speaker
Elizabeth Hume-O'Haire, Ohio State University, discussant

Phonological change through the interaction of L1-L2 in bilinguals and language learners
James Flege, University of Alabama, invited speaker
Norma Mendoza-Denton, University of Arizona, discussant

University of Illinois at Urbana-Champaign | Department of Linguistics | Beckman Institute for Advanced Science & Technology | National Science Foundation

LabPhon 9
Change in Phonology

Ninth Conference on Laboratory Phonology
Thursday 24 June - Saturday 26 June, 2004

LabPhon 9 hosts

Location

Important Dates

Program Committee

Conference Grants

Downloads

- **Schedule [pdf]**
- **Shuttle Service**
- **Style Guidelines [pdf]**

Please follow this link to the visit the 10th Conference on Laboratory Phonology page: *Variation, Detail and Representation*

Welcome to the 9th Conference on Laboratory Phonology

LabPhon is an international forum for interdisciplinary research on the sound structure of languages. Participants in LabPhon seek to communicate across the traditional boundaries that separate phonology, as a branch of theoretical linguistics, from the study of speech production, speech perception, spoken language acquisition, computer speech processing, and other disciplines concerned with human speech.

The theme for LabPhon 9, ***Change in Phonology***, addresses questions related to the evolution of language within a speech community and the development of language within the individual speaker/hearer, and includes issues of broad social interest, such as child language development, language pedagogy, and technologies for human-computer interaction.

Themes and Invited Participants

Acquisition as change: L1 phonology
LouAnn Gerken, University of Arizona , invited speaker
Stefan Frisch, University of South Florida, discussant

Phonological models of variation in computer speech processing
Julia Hirschberg, Columbia University, invited speaker
Carol Espy-Wilson, University of Maryland, discussant

Prosodic influence on change in sound patterns
Cécile Fougeron, CNRS-Université Paris 3, invited speaker
Kenneth de Jong, Indiana University, discussant

Social factors in phonetic variation
Gerry Docherty, University of Newcastle upon Tyne, invited speaker
Henrietta Jonas-Cedergren, Université du Québec à Montréal, discussant

Mechanisms of sound change
Jonathan Harrington, University of Kiel, invited speaker
Elizabeth Hume-O'Haire, Ohio State University, discussant

Phonological change through the interaction of L1-L2 in bilinguals and language learners
James Flege, University of Alabama, invited speaker
Norma Mendoza-Denton, University of Arizona, discussant

University of Illinois at Urbana-Champaign | Department of Linguistics | Beckman Institute for Advanced Science & Technology | National Science Foundation

LabPhon 9
Change in Phonology

Ninth Conference on Laboratory Phonology
Thursday 24 June - Saturday 26 June, 2004

Call for Papers

The call for papers deadline has passed.

Please see the Program link for the program of oral and poster presentations.

University of Illinois at Urbana-Champaign
Department of Linguistics
Beckman Institute for Advanced Science & Technology
National Science Foundation

LabPhon 9
Change in Phonology

Ninth Conference on Laboratory Phonology
Thursday 24 June - Saturday 26 June, 2004

Program

Oral Presentations

- 24th June, 2004
- 25th June, 2004
- 26th June, 2004

Poster Presentations

24th June, 2004

7:30	Continental Breakfast	
8:25	Welcome Address	Pierre Wiltzius, Director, Beckman Institute
8:30	Introductions	Jennifer Cole, José I. Hualde
Session 1: Acquisition as change: L1 phonology		Chair: Cynthia Fisher (UIUC)
8:45	LouAnn Gerken (University of Arizona) *invited speaker* with Tania Zamuner (University of Nijmegen)	Exploring the Basis for Generalization in Language Acquisition
9:45	Angela Grimm (University of Groningen)	The Prosodic Pattern of Words and Phrases in the Acquisition of German
10:15	Coffee Break	
10:30	Mary E. Beckman, Benjamin Munson, & Jan Edwards (Ohio State University and University of Minnesota)	Vocabulary growth and developmental expansion of types of phonological knowledge
11:00	Sharon Peperkamp & Emmanuel Dupoux (Laboratoire de Sciences Cognitives et Psycholinguistique, Paris and Université de Paris 8)	The acquisition of abstract phoneme categories
11:30	Stefan Frisch (University of South Florida)	*discussant*
12:00-1:00	Lunch break: *Catered lunch on-site*	
Session 2: Prosodic influence on change in sound patterns		Chair: Chilin Shih (UIUC)
1:00	Cécile Fougeron (CNRS-U Paris III) *invited speaker*	Resyllabification revisited: the case of enchaînement in French
2:00	Mariapaola D'Imperio, Hélène Lœvenbruck, Caroline Menezes, Noël Nguyen & Pauline Welby (Laboratoire Parole et Langage, Aix-en-Provence and Institut de la Communication Parlée, Grenoble)	Are tones aligned with articulatory events? Evidence from Italian and French
2:30	Coffee Break	
2:45	Gina Garding & Amalia Arvaniti (University of California, San Diego)	Dialectal variation in the rising accents of American English
3:15	Laura Dilley (MIT and Harvard University)	Fundamental frequency extrema on

	Cancelled	weak syllables affect the relative prominence of strong syllables
3:15	Kenneth de Jong (Indiana University)	*discussant*

Session 3: Non-thematic		Chair: Erin O'Rourke (UIUC)
3:45	Amanda Miller-Ockhuizen, Levi Namaseb & Khalil Iskarous (Cornell University, University of Namibia and Haskins Laboratory)	Posterior constriction location differences in Click types
4:15	Mariko Sugahara & Alice Turk (University of Edinburgh)	Phonetic Reflexes of Morphological Boundaries Under Different Speech Rates

5:00-6:30	Poster session 1 Wine & Cheese Reception *Beckman Institute East Atrium*

25th June, 2004

7:30	Continental Breakfast

Session 4: Social factors in phonetic variation		Chair: Zsuzsanna Fagyal (UIUC)
8:30	Gerry Docherty (University of Newcastle upon Tyne) *invited speaker*	Speech in its natural environment: accounting for social factors in phonetic variability
9:30	Jane Stuart-Smith (University of Glasgow)	Empirical evidence for gendered speech production: /s/ in Glaswegian
10:00	Coffee Break	
10:15	Elliott Moreton & Erik Thomas (University of North Carolina at Chapel Hill and North Carolina State University)	Origins of Canadian Raising in Voiceless-Coda Effects: A Case Study in Phonologization
10:45	Paul Warren, Jen Hay & Brynmor Thomas (Victoria University of Wellington and University of Canterbury)	The loci of sound change effects in recognition and perception
11:15	Henrietta Jonas-Cedergren, Université du Québec à Montréal	*discussant*

11:45-1:00	Lunch break: *Catered lunch on-site*

Session 5: Mechanisms of sound change		Chair: Hans Henrich Hock (UIUC)
1:00	Jonathan Harrington (University of Kiel) *invited speaker*	Evidence for a relationship between synchronic variability and diachronic change in the Queen's annual Christmas broadcasts
2:00	Susan G. Guion & Ratree P. Wayland (University of Oregon and University of Florida at Gainesville)	Aerodynamics of [r] in tonogenesis
2:30	Coffee Break	
2:45	John Hajek & Mary Stevens (University of Melbourne)	Mechanisms of sound change in Romance: From gemination to degemination in Italy
3:15	Carlos Gussenhoven (University of Nijmegen)	A vowel height split explained: Compensatory listening and speaker control
3:45	Elizabeth Hume (Ohio State University)	*discussant*

4:15-4:45	General Discussion

5:00-6:30	Poster session 2 *Beckman Institute East Atrium*

5:30	Wine & Appetizers	
6:30	Banquet dinner *Beckman Institute West Atrium*	

26th June, 2004

7:30	Continental Breakfast	
Session 6: Phonological models of variation in computer speech processing		Chair: Mark Hasegawa-Johnson (UIUC)
8:30	Julia Hirschberg (Columbia University) *invited speaker* with Agus Gravano (Columbia University), Ani Nenkova (Columbia University), Elisa Sneed (Northwestern University) & Gregory Ward (Northwestern University)	Intonational overload: Uses of the H* !H* L- L% contour in read and spontaneous speech
9:30	Robert Kirchner (University of Alberta)	Exemplar-based phonology and the time problem: a new representational technique
10:00	Coffee Break	
10:15	Hosung Nam (Yale University and Haskins Laboratory)	A Competitive, Coupled Oscillator Model of Moraic Structure
10:45	Carol Espy-Wilson (University of Maryland)	*discussant*

Session 7: Non-thematic		Chair: Gary Dell (UIUC)
11:15	Bob McMurray & David Gow (University of Rochester)	Tracking the timecourse of multiple context effects in assimilated speech
11:45	Benjamin Munson (University of Minnesota)	Lexical Access and Hyperarticulation
12:15-1:15	Lunch break: *Catered lunch on-site*	

Session 8: Phonological change through the interaction of L1-L2 in bilinguals and language learners		Chair: Rajka Smiljanic (Northwestern)
1:15	James Flege (University of Alabama) *invited speaker*	Language Contact in Bilingualism: Phonetic System Interactions
2:15	Isabelle Darcy, Sharon Peperkamp & Emmanuel Dupoux (Laboratoire de Sciences Cognitives et Psycholinguistique, Paris and Université de Paris 8)	Perceptual learning and plasticity in L2 learners: building a new system for phonological processes
2:45	Mariko Yanagawa (Yale University)	Consonant Timing in L2 English: The Emergence of a Default Pattern
3:15	Coffee Break	
3:30	Ghada Khattab (University of Newcastle)	Variation in vowel production by English-Arabic bilinguals
4:00	Norma Mendoza-Denton (University of Arizona)	*discussant*
4:30-5:00	General Discussion	

LabPhon 9
Change in Phonology

Ninth Conference on Laboratory Phonology
Thursday 24 June - Saturday 26 June, 2004

US Minority Students Participation Grant

International Scholar Participation Grant

University of Illinois at Urbana-Champaign | Department of Linguistics | Beckman Institute for Advanced Science & Technology

LabPhon 9
Change in Phonology

Ninth Conference on Laboratory Phonology
Thursday 24 June - Saturday 26 June, 2004

Abstracts

Oral Presentations
Poster Presentations

University of Illinois at Urbana-Champaign
Department of Linguistics
Beckman Institute for Advanced Science & Technology

LabPhon 9
Change in Phonology

Ninth Conference on Laboratory Phonology
Thursday 24 June - Saturday 26 June, 2004

Contact Information

E-mail	Please mail labphon9@uiuc.edu with any questions.
Postal Address	**LabPhon 9** Department of Linguistics, 4080 Foreign Languages Building, 707 South Matthews Avenue, Urbana, IL 61801
Phone & Fax	Phone: (217) 333-3563 Fax: (217) 333-3466

University of Illinois at Urbana-Champaign | Department of Linguistics | Beckman Institute for Advanced Science & Technology

LabPhon 9
Change in Phonology

Ninth Conference on Laboratory Phonology
Thursday 24 June - Saturday 26 June, 2004

LabPhon 9 Hosts

Department of Linguistics

Beckman Institute for Advanced Science & Technology

University of Illinois at Urbana-Champaign

Department of Linguistics

Beckman Institute for Advanced Science & Technology

LabPhon 9
Change in Phonology

Ninth Conference on Laboratory Phonology
Thursday 24 June - Saturday 26 June, 2004

Location

Beckman Institue for Advanced Science & Technology
University of Illinois at Urbana-Champaign
405 North Mathews Avenue
Urbana, IL 61801 USA

Telephone: (217) 244-1176
Fax: (217) 244-8371

University of Illinois at Urbana-Champaign | Department of Linguistics | Beckman Institute for Advanced Science & Technology | National Science Foundation

LabPhon 9
Change in Phonology

Ninth Conference on Laboratory Phonology
Thursday 24 June - Saturday 26 June, 2004

Important Dates

Abstract submission: October 31, 2003
Notification of acceptance: January 15, 2004
Submission of accepted papers: April 15, 2004

University of Illinois at Urbana-Champaign | Department of Linguistics | Beckman Institute for Advanced Science & Technology | National Science Foundation

LabPhon 9
Change in Phonology

Ninth Conference on Laboratory Phonology
Thursday 24 June - Saturday 26 June, 2004

Program Committee

Jennifer S. Cole (Linguistics), Chair
José I. Hualde (Linguistics; Spanish, Italian & Portuguese), Co-Chair
Gary Dell (Psychology)
Zsuzsanna Fagyal (French)
Cindy Fisher (Psychology)
Mark Hasegawa-Johnson (Electrical & Computer Engineering)
Stephen Levinson (Electrical & Computer Engineering)
Daniel Silverman (Linguistics)

University of Illinois at Urbana-Champaign
Department of Linguistics
Beckman Institute for Advanced Science & Technology
National Science Foundation

LabPhon 9
Change in Phonology

Ninth Conference on Laboratory Phonology
Thursday 24 June - Saturday 26 June, 2004

[Print pdf version]

Conference participants are invited to a Welcome Reception on Wednesday, **June 23, 6-8 p.m**. at the **Beckman Institute West Atrium**. The updated conference program and other information will be available at the reception, and on-site registration will be processed for those who have not pre-registered. Wine, soft drinks and hors d'oeuvres will be served.

University of Illinois at Urbana-Champaign | Department of Linguistics | Beckman Institute for Advanced Science & Technology | National Science Foundation

9th Conference on Laboratory Phonology
:: University of Illinois at Urbana-Champaign ::
:: 24th -26th June, 2004 ::

Wednesday, June 23 :: Welcome Reception :: 6:00 p.m. :: *Beckman Institute West Atrium*

Day 1: Thursday, June 24

7:30 a.m.	CONTINENTAL BREAKFAST
8:25 a.m.	Welcome Address, Pierre Wiltzius, Director, Beckman Institute
8:30 a.m.	Introductions by Jennifer Cole and José I. Hualde
8:45 a.m.– 10:15 a.m.	**Session 1: Acquisition as change: L1 phonology**
10:15 a.m. – 10:30 a.m.	COFFEE BREAK
10:30 a.m. – 12:00 noon	**Session 1: Acquisition as change: L1 phonology (*Continued*)**
12:00 noon - 1:00 p.m.	LUNCH BREAK *Catered lunch on-site*
1:00 p.m. – 2:30 p.m.	**Session 2: Prosodic influence on change in sound patterns**
2:30 p.m. – 2:45 p.m.	COFFEE BREAK
2:45 p.m. - 3:45 p.m	**Session 2: Prosodic influence on change in sound patterns (*Continued*)**
3:45 p.m. – 4:45 p.m.	**Session 3: Non-thematic**
5:00 p.m. - 6:30 p.m.	POSTER SESSION I *Beckman Institute East Atrium*
5:00 p.m. - 6:30 p.m.	RECEPTION - Appetizers and beverage service

Day 2: Friday, June 25

7:30 a.m.	CONTINENTAL BREAKFAST
8:30 a.m. – 10:00 a.m	**Session 4: Social factors in phonetic variation**
10:00 a.m. – 10:15 a.m.	COFFEE BREAK
10:15 a.m. – 11:45 a.m.	**Session 4: Social factors in phonetic variation (*Continued*)**
11:45a.m - 1:00 p.m.	LUNCH BREAK *Catered lunch on-site*
1:00 p.m. – 2:30 p.m.	**Session 5: Mechanisms of sound change**
2:30 p.m. – 2:45 p.m.	COFFEE BREAK
2:45 p.m. – 4:15 p.m.	**Session 5: Mechanisms of sound change (*Continued*)**
4:15 p.m. - 4:45 p.m.	General Discussion
5:00 p.m. – 6:30 p.m.	POSTER SESSION II *Beckman Institute East Atrium*
5:30 p.m.	RECEPTION - Appetizers and beverage service
6:30 p.m.	BANQUET DINNER *Beckman Institute West Atrium*

Day 3: Saturday, June 26

7:30 a.m.	CONTINENTAL BREAKFAST
8:30 a.m. – 10:00 a.m.	**Session 6: Phonological models of variation in computer speech processing**
10:00 a.m. – 10:15 a.m.	COFFEE BREAK
10:15 a.m. – 11:15 a.m.	**Session 6: Phonological models of variation in computer speech processing (*Continued*)**
11:15 a.m. – 12:15 p.m.	**Session 7: Non-thematic**
12:15 p.m. - 1:15 p.m.	LUNCH BREAK *Catered lunch on-site*
1:15 p.m. – 3:15 p.m.	**Session 8: Phonological change through the interaction of L1-L2 in bilinguals and language learners**
3:15 p.m. – 3:30 p.m.	COFFEE BREAK
3:30 p.m. – 4:30 p.m.	**Session 8: Phonological change through the interaction of L1-L2 in bilinguals and language learners (*Continued*)**
4:30 p.m. – 5:00 p.m.	General Discussion

LabPhon 9
Change in Phonology

Ninth Conference on Laboratory Phonology
Thursday 24 June - Saturday 26 June, 2004

Conference Shuttle Service

Participants staying at the Hawthorne Suites and Chancellor hotels can take advantage of a shuttle service provided by the LabPhon 9 organizers for transportation between the hotel and the conference site. Shuttle vans will be identified with a LabPhon 9 sign on the dashboard, and transportation will be provided at the following times:

Wednesday: 5:30 p.m - 6:00 p.m. from Hawthorne Suites and Chancellor to Beckman Institute

8:00 p.m - 8:30 p.m. between Beckman Institute and Hawthorne Suites and Chancellor

Thursday: 7:30 a.m. - 8:15 a.m. from Hawthorne Suites and Chancellor to Beckman Institute

6:30 p.m. - 8:00 p.m. between Beckman Institute and Hawthorne Suites and Chancellor

Friday: 7:30 a.m. - 8:15 a.m. from Hawthorne Suites and Chancellor to Beckman Institute

8:00 p.m. - 9:00 p.m. between Beckman Institute and Hawthorne Suites and Chancellor

Saturday: 7:30 a.m. - 8:15 a.m. from Hawthorne Suites and Chancellor to Beckman Institute

5:00 p.m. - 5:30 p.m. between Beckman Institute and Hawthorne Suites and Chancellor

Evening shuttle service between Beckman Institute, Hawthorne Suites and the Chancellor Hotel will include a stop in Downtown Champaign near a variety of restaurants, clubs and bars.

University of Illinois at Urbana-Champaign
Department of Linguistics
Beckman Institute for Advanced Science & Technology
National Science Foundation

Guidelines for paper submissions

The LabPhon 9 volume will be published by Mouton de Gruyter. Your paper will be formatted by the publisher, and some important publisher's style guidelines are provided below. **Your paper should not exceed the length limit specified below,** allowing an additional 200 words for your abstract. Beyond the word limit, you are also allowed a maximum of two pages for tables and figures. If you need more space for tables and figures you are advised to shorten the text portion of your paper accordingly. Given the number of papers that we aim to include, it will be very important that contributors stay within the length guidelines. We will not be able to accommodate longer papers. Please address questions to the editors of the LabPhon 9 volume (Jennifer Cole and José I. Hualde).

Length limits

Invited papers: **11,000** words
Regular papers: **8,000** words
Discussant papers: **4,000** words

Please conform to IPA standards for phonetic transcription, and use the SIL IPA93/ Doulos phonetic fonts (from the Encore family of fonts), which you can download for free from http://www.sil.org/computing/fonts/encore-ipa.html. In addition to the abbreviation guidelines from the publisher's style sheet, the editors ask that you adhere to the following abbreviations for technical terms:

F_0	fundamental frequency
F1, F2, etc.	first formant, second formant, etc.
ms.	milliseconds
s.	seconds
P_s	subglottal pressure
Hz	Hertz
dB	decibel
p	significance level
Fig.	figure

Publisher's style sheet

General points

Most of the formatting will be done either by typesetters or by a template that converts your paper into the Mouton style (then the editor of the book will take care of this). Please support the editor(s) by considering the following guidelines while preparing your paper.

Type area & type sizes

You do not have to specify a type area. But please note that the final manuscript will be restricted to a

- Width of 118 mm (4.65 inches)
- Length of 180 mm (7.09 inches)

Please avoid illustrations, figures and examples that would exceed these measures (in Times, 10 Point). You don't have to specify fonts and type sizes. Your document will later be converted to Times, 11 Point (and Times 10 Point for examples).

Headings and line spacing

– Do not give the academic affiliation of the author of an article.
– Do not put a period at the end of a heading.
– Capitalize only the first letter of the first word and of proper nouns and adjectives, e.g., ”The capitalization of titles in English” not ”The Capitalization of Titles in English”. Please also capitalize the first letter of the first word after a colon.
– Aim to use no more than 3 levels of heading. However, if a fourth-grade heading is required, use the same format (i.e., 11 pt italic) as for the third-grade heading.

Orthography
– Both American and British English forms are acceptable, but spelling and punctuation must be consistent throughout.
– Please note that in British English the *-ize* ending should be used in preference to *–ise* where both spellings are in use (e.g., *criticize*, *recognize*).

Quotations
– Short quotations (fewer than 60 words) should be run on in text and enclosed in double quotation marks. Single quotation marks enclose quotations within quotations.
– Longer quotations should appear as a separate block.
– All quotations in languages other than English should be followed by the translation in square brackets.
– Omissions are indicated by ellipsis points without brackets.
– Any insertions by the author are to be enclosed in square brackets: [emphasis mine].

Citations
– Full bibliographical details are given in the reference section at the end of the book or article.
– Brief citations are used within the text:

One author:	(Bouissac 1985)
Two authors:	(Smith and Jones 1995)
Three authors:	(Uexküll, Geigges, and Host 1993)
Four or more authors:	(Smith et al. 1990: 38)
(but list all authors in the reference entry)	
Works by one author:	(Bouissac 1987a, 1987b, 1994)
Works by different authors:	(Bouissac 1985; Deakin 1993)
No dropped digits in inclusive page numbers:	(Hockett 1964: 140–145)
Volume number:	(Balat and Dove 1992, 1: 210)
Reprints:	(Dickens [1854] 1987: 73)

Only give the original date at the first mention, in all subsequent citations give reprint date only: (Dickens 1987: 73). In citations of reprints of more recently published titles, only give the date of the reprint in the citation (full details should only be given in the reference section, see section 15).
– The date is always given in brackets: ”Bloomfield (1933: 123–125) introduced the term . . .”; ”In his (1922) article Sapir argued that . . . ”
– Give page numbers in full: do not use ”f.”, ”ff.”.
– Avoid referring to a whole book: give exact page numbers whenever possible. Always give the page numbers for quotations.
– Always give the full author-date citation: do not use ”op. cit.”, ”loc. cit.” or ”ibid.”.

Typeface, emphasis, and punctuation
Italics should be used for:
– words, phrases, and sentences treated as linguistic examples.
– foreign-language expressions.
– titles of books, published documents, newspapers, and journals.
Italics may also be used for*:*
– drawing attention to key terms in a discussion at first mention only. Thereafter, these terms should be set

in roman. However, please keep the use of italics to a minimum.
– emphasizing a word or phrase in a quotation indicating [emphasis mine].

Bold may be used sparingly to draw attention to a particular linguistic feature in numbered examples (not in running text).

Underlining or *capital letters* should not be used for emphasis.

Quotation marks:
– Single quotation marks should be used for the translation of non-English words, e.g., *cogito* 'I think'.
– Double quotation marks should be used in all other cases, i.e. direct quotations in running text.
– Use rounded quotation marks (". . .") not "straight" ones.

Dashes:
– spaced EN dashes are used as parenthetical dashes. An EN dash is longer than a hyphen, "word – word". Do not use double hyphens. Standard WORD shortcut: CTRL + - (hyphen key).
– unspaced EN dashes should be used between inclusive page numbers, 153–159.
– please use EN dashes (not bullets) for lists without numbering.

Periods should be placed before superscript note numbers e.g., . . .word.[7]

Spacing: Type one space (not two) after periods, commas and colons.

Brackets: Do not use double round brackets: brackets within brackets should be square brackets.

Abbreviations
– Use only the most common abbreviations.
– Periods should be used after abbreviations, but not in acronyms:
Common abbreviations: cf. Dr. ed. eds. e.g. et al. etc. i.e. no. trans. vol. vols.
Example acronyms: USA ICLA ELT
– Initials require periods and have a space between them, e.g. Ronald W. Langacker.
– Abbreviations common in linguistics (NP, V, ACC) may be used in numbered examples, but the terms should be written out in full in the text.
– Write out names of theories, titles of books or names of publishers: "the Spatialization of Form Hypothesis", not "the SFH"; "Oxford University Press", not "OUP."

Examples
– Number examples chapter by chapter in authored works, and article by article in edited works.
– Foreign-language examples should be presented as follows. Use tabs (not the space bar) to align examples and glosses. Please also use italics for examples in English, see (2b):

(1) *Original language in italics.*
Word-for-word gloss in roman (plain) type with correct alignment.
'Idiomatic translation in single quotation marks.'

(2) a. *Mampianatra anglisy an-d Rabe aho.*
Cause-learn English ACC-Rabe I
'I am teaching Rabe English.'

b. *I sent the artefacts to an anthropologist.*

References to examples in the text should take the form "see (1a) and (1b)" with both number and letter in

brackets.

Tables, figures, and illustrations
– If there are figures to be inserted, please have them professionally drawn and send us the originals, not photocopies.
– If photographs are to be inserted, the prints must be provided.
– Tables and figures should be numbered consecutively throughout an authored work (only number article by article in an edited work).
– Title of a table is to be placed flush left above the table.
– Title of a figure is to be placed flush left below the figure.
– In tables use the full width of the type area and a minimum of two fine horizontal lines (no vertical lines).
– Avoid using tints as this can affect legibility.
Example table:

Table 1. Infinitive suffixes [Captions of more than one line will later automatically be formatted with a hanging indent.]

for to	*to*	*a*	*to*-forms in %	other forms in %
Rolls 2	3	12	80	20
Rolls 3	2	4	1 57	43
Rolls 4a	2	30	94	6
Rolls 4b	2	12	86	14

Maps
If your book or article treats a language that is little known, or if it will help the reader to identify geographic locations, it may be useful to include a map. Please bear in mind the following:
– All maps should be enclosed in an outer frame.
– All labels must remain legible when the maps are reduced to fit the page.
– Avoid using tints as this can affect legibility.
– All national boundaries in the area shown on the map must be drawn in and countries, counties or federal states bordering on the area of main interest must be named, i.e., a land-locked area should not be represented as if it were an island.

Appendixes and notes
– Location: an appendix is placed at the end of the text, before the notes. The notes section should be placed at the end of an article in an edited work or at the end of an authored work, before the reference section.
– In an authored work, appendixes and notes both start on a new page (right- or left-hand). The heading should be treated as a chapter title. In an article in an edited work, the notes follow directly at the end of the article before the references, with a first-grade heading.
– Notes should be numbered consecutively throughout the text.
– Note numbers in the text should be superscript (small raised) numbers[1] without punctuation or brackets.
– The note number should directly follow the word in question or a punctuation mark, with no blank space.[2]

References
– The reference section immediately follows the notes. In an authored work (but not in an article in an edited work) the section starts on a new page (right- or left-hand).
– The reference section should contain all works cited and only those.
– All works must be listed in alphabetical order.
– Whenever possible give the full first names of authors and editors.
– Initials require periods and have a space between them: Ronald W. Langacker.
– Arrange all works under a particular author’s/editor’s name in chronological order.
– Give the full title and subtitle of each work.
– Titles of published books and journals are capitalized and italicized. Unpublished works, such as Ph.D.

dissertations, and the titles of articles in journals or edited works are neither capitalized nor italicized.
– Give the inclusive page numbers of articles in journals or edited works. Do not drop digits in inclusive page numbers.
– Give both the place of publication and the name of the publisher.
– Do not use abbreviations for names of journals, book series, publishers or conferences.
– Do not use ”et al.” but list all author/editor names.
– Translate titles in languages other than French, German, Spanish and Italian into English. The translation should be set in roman and placed in square brackets following the italicized original title.
– Use tabs, not the space bar, to position dates of publication and titles.

The References do not have to be formatted – just use tabs so that there are three fields:
Name Year Title etc
and make use of italics as shown in the example entries below

Example entries:

Book (authored work):
Chomsky, Noam 1986 *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger Publishers.

Edited work:
Preston, Dennis R. (ed.) 1999 *Handbook of Perceptual Dialectology*. Vol. 1. Amsterdam/Philadelphia: Benjamins.

Article in an edited work:
Heller, Monica 2001 Gender and public space in a bilingual school. In *Multilingualism, Second Language Learning, and Gender*, Aneta Pavlenko, Adrian Blackledge, Ingrid Piller, and Marya Teutsch-Dwyer (eds.), 257–282. (Language, Power and Social Process 6.) Berlin/New York: Mouton de Gruyter.

Journal article:
Neuman, Yair, Yotam Lurie, and Michele Rosenthal 2001 A watermelon without seeds: A case study in rhetorical rationality. *Text* 21 (4): 543–565.

Reprint:
Jakobson, Roman, and Morris Halle 2002 Reprint. *Fundamentals of Language*. 2d ed. Berlin/New York: Mouton de Gruyter. Original edition, The Hague: Mouton, 1956.

Unpublished dissertation:
Kim, Yong-Jin 1990 Register variation in Korean: A corpus-based study. Ph.D. diss., Department of Linguistics, University of Southern Carolina.

Translated title:
Haga, Yasushi 1998 *Nihongo no Shakai Shinri* [Social psychology in the Japanese language]. Tokyo: Ningen no Kagaku Sha.

Several works by one author/editor:
Vennemann, Theo (ed.) 1989 *The New Sound of Indo-European: Essays in Phonological Reconstruction.* (Trends in Linguistics: Studies and Monographs 41.) Berlin/New York: Mouton de Gruyter.
2000a From quantity to syllable cuts: On so-called lengthening in the Germanic Languages. *Journal of Italian Linguistics/Rivista di Linguistica* 12: 251–282.

2000b Triple-cluster reduction in Germanic: Etymology without sound laws? *Historische Sprachwissenschaft* [Historical linguistics] 113: 239–258.

Index
The index should be compiled by the volume editors. Contributors should provide them with a list of index entries for their paper.

How to do *x*?
Please do not spend too much time trying to find a solution which may be unacceptable either from a typographic point of view, or because of particular formal requirements of the series your book will be published in. Please feel free to contact us at any time, and we will be more than happy to advise you.

Our address: Mouton de Gruyter
Genthiner Strasse 13
10785 Berlin
Germany

Phone: +49 30 26005 235
Fax: +49 30 26005 351
Email: mouton.editorial@degruyter.de

LabPhon 9
Change in Phonology

Ninth Conference on Laboratory Phonology
Thursday 24 June - Saturday 26 June, 2004

Instructions for poster presenters:

Posters will be mounted on display boards and must be no larger than 45 inches (height) x 58 inches (width). A small number of wider display boards (93 inches) are available by request. Tacks for mounting posters will be provided.

Poster Sessions

- Session I: **Thursday 24th June, 2004**
- Session II: **Friday 25th June, 2004**

5:00-6:30 **Thursday 24th June, 2004**	***Beckman Institute East Atrium***
Julie Carson-Berndsen, Anja Geumann & Moritz Neugebauer (University College Dublin)	Embracing Multilinguality: Defining Phonetic Features for Speech Technology
Ivan Chow (University of Toronto)	A study of the melodic contour of Mandarin and French through Syntactic Structural Ambiguity
Daniel A. Dinnsen (Indiana University)	On the emergence and loss of opacity effects in acquisition
Judith A. Gierut (Indiana University)	Something from Nothing: Permutability of Competence and Performance in Phonological Acquisition
Christine Haunz (University of Edinburgh)	Perceived similarity: universal or language-dependent?
Claudia Kuzla & Taehong Cho (Max Planck Institute for Psycholinguistics)	Prosodic and phonotactic influences on fricative voicing assimilation in German
Eleonora C. Albano (Universidade Estadual de Campinas)	Two Complementary Phonotactic Constraints Interfacing Grammar with Phonetics
Z.S. Bond, Dace Markus & Verna Stockmal (Ohio University and University of Latvia)	Sixty years of bilingualism affect the pronunciation of Latvian vowels
Josefina Carrera-Sabaté (Universitat de Barcelona and Universitat de Lleida)	Instruction and Phonetic Change
Véronique Delvaux, Didier Demolin & Alain Soquet (Université de Mons-Hainaut, Universidade de Sao Paolo and Université Libre de Bruxelles)	Sound change as a result of mimetic interactions between speakers: Implementation of imitation in the laboratory
Haruo Kubozono, Nobuko Kibe & Ichiro Ota (Kobe University and Kagoshima University)	Accent Changes in an Endangered Japanese Dialect
Tivoli Majors (University of Missouri)	Phonetic variation in Missouri: The low vowel merger in transition
Andrea Pearman (Universitat Autònoma de Barcelona)	Sound change in the learner: The perception of connected speech
Maria-Josep Sole & Andrea Pearman (Universitat Autònoma de Barcelona)	Effects of syllable position on sound change: Aerodynamic and perceptual data on final ficative weakening
Ioana Chitoran (Dartmouth College)	The effect of pitch accent and word position on the production of vowel sequences: a comparison of Spanish, Romanian, and French
Natasha Warner, Lynnika Butler & Takayuki Arai (University of Arizona and Sophia University)	Intonation as a speech segmentation cue: Effects of speech style
Kimiko Tsukada, Denis Burnham, Sudaporn Luksaneeyanawin, Niratasai Krachaikiat & Sorabud	The effect of tone on vowel duration in Thai: A

Rungrojsuwan	developmental study
Ewa Jacewicz, Joseph C. Salmons & Robert A. Fox (Ohio State University and University of Wisconsin)	Prosodic domain effects and vocalic chain shifts
Kenji Yoshida (Shoin University)	The delay of the timing of F0 minima in Akita Japanese
Ivan Yuen (University of Edinburgh)	Normalisation of downtrend --- local or global effect?
Shakuntala Mahanta & K.G. Vijayakrishnan (Utrecht Institute of Linguistics and Central Institute of English and Foreign Languages-Hyderabad)	Linking of Pitch to Mora: An Experimental Approach
Heike Lehnert-LeHouillier (University at Buffalo and Haskins Laboratories)	The Separation of Pitch and Stress in Onondaga
Bert Remijsen (Leiden University)	Segmental and prosodic factors in a Matbat vowel change
Yohann Meynadier (University of Edinburgh)	Gradual linguopalatal variations due to a 4-level prosodic hierarchy in French
Cynthia G. Clopper (Indiana University)	Towards an Exemplar-Based Phonological Model of the Perception of Dialect Variation
Katherine Crosswhite & Michael K. Tanenhaus (University of Rochester)	Online Use of Phonetic Detail in Spoken Word Recognition
Michelle L. Gregory, Julie Sedivy, Anjula Joshi & Danial Grodner (University at Buffalo, Brown University and Univeristy of British Columbia)	Isolating the cognitive processes that underlie disfluencies and phonological form: Evidence from an on-line production experiment
Michele L. Morrisette (Indiana University)	Recursive patterns of phonological change in the lexicon
Khalil Iskarous, Heike Lehnert, Louis Goldstein & Mark Tiede (Haskins Laboratories)	Distinguishing Speaker and Linguistic Variation Using Configurable Articulatory Synthesis
Didier Demolin (Universidade de Sao Paulo and Université Libre de Bruxelles)	Changes in articulatory settings and aerodynamic conditions as sources of sound change
Caroline L. Smith (University of New Mexico)	Discourse-level effects on durations in synthesized English and French: testing listener preferences
Marie-Hélène Côté & Geoffrey Morrison (Université d'Ottawa and University of Alberta)	Experimental evidence and the nature of the schwa/zero alternation in French
Paul M. De Decker (New York University)	Co-articulation and Speech Acoustics in Phonological Change: some experimental data
Manuel Díaz-Campos & Richard J File-Muriel (Indiana University)	The effect of phonetic factors in the perception of phonological variation: An experimental study of internal constraints of syllable-final /s/
T. A. Hall, Silke Hamann & Marzena Zygis (University of Leipzig and ZAS)	The phonetic motivation of stop assibilation
Alexei Kochetov (Simon Fraser University and Haskins Laboratories)	From phonetic differences to phonological asymmetries: Secondary articulation contrasts in liquids
Tomoko Kozasa (University of Hawaii at Manoa)	Sustaining phonemic distinctions: A study of phonetic cues in Japanese long vowels
Tae-Jin Yoon (University of Illinois at Urbana-Champaign)	Asymmetry in Laryngeal Metathesis
Scott Myers (University of Texas at Austin)	Vowel duration in Kinyarwanda: Effects of quantity, height and position
	Flatly contradicted: The non-

	Ryan Shosted (University of California-Berkeley)	equivalency of a distinctive feature in Hindi and Levantine Arabic
	Katherine S. White, Cecilia Kirk, & James L. Morgan, Sharon Peperkamp, & Emannuel Dupoux (Brown University, Laboratoire de Sciences Cognitives et Psycholinguistique, EHESS - ENS – CNRS, Paris and Université de Paris 8)	Distributional and statistical bases of allophonic groupings
	Man Gao (Yale University)	Gestural Analysis on Mandarin Tongue Twisters
	Matthew Goldrick & Sheila Blumstein (Brown University)	The Bases of Speech Errors: Local and Non-Local Phonetic Traces
5:00-6:30	**Friday 25th June, 2004**	***Beckman Institute East Atrium***
	Alan Bell, Jason Brenier, Michelle Gregory, Dan Jurafsky & Cynthia Girand (University of Colorado-Boulder, State University of New York-Buffalo and Stanford University)	Ranges and levels of predictability effects on word durations in conversational English
	Melissa A. Redford & Risto Miikkulainen (University of Oregon and University of Texas)	Phonological Differences Emerging from Slow versus Fast Lexicon Expansion
	Della Chambless (University of Massachusetts)	Asymmetries in Intermediate Stages of Cluster Acquisition
	Chun-Mei Chen (University of Texas)	Phonological change of Taiwan Mandarin -Evidence from the Three Presidential Candidates' Speech
	Taehong Cho & James McQueen (Max Planck Institute for Psycholinguistics)	Perceptual reality of phonetically-driven phonology: Place assimilation and consonant cluster simplification at different prosodic boundaries
	Paola Escudero & Paul Boersma (Utrecht University and University of Amsterdam)	L2 re-categorization of an 'old' phonological contrast
	Masako Hirotani (University of Massachusetts)	Does Prosodic Phrasing Correlate with Wh-scope in Japanese?
	Byung-jin Lim (University at Albany, State University of New York)	Effects of L1 phonology and orthography on L2 perception: The use of Korean Hangul in perception of English voicing contrasts
	Bruce Morén (University of Tromsø)	The Phonetics and Phonology of Front Vowels in Staten Island: when the traditional descriptions and the facts do not agree
	Geoffrey Stewart Morrison (University of Alberta)	Are functional constraints active synchronically?
	Ana Sánchez Muñoz (University of Southern California)	Convergence or Divergence? Accommodation in the Vocalic System of Texan English
	James Myers & Jane Tsay (National Chung Cheng University)	Exploring performance-based predictors of phonological judgments in Mandarin
	Pawel M. Nowak (University of California-Berkeley)	Polish sibilants: studying perceptual cues to explain the directionality of sound change
	Miren L. Oñederra (UPV-EHU, University of the Basque Country)	The importance of social factors in the loss or recovery of phonological units
	Michela Russo (Univ. de Paris 8/UMR 7023-CNRS)	Articulatory weakening of an apical obstruent: lambdacized and rhotacized forms in Italian dialects
	Caroline Wiltshire & James D. Harnsberger (University of Florida)	The influence of Gujarati and Tamil L1s on Indian English
	Zheng Xu (State University of New York-Stony	The interlanguage phonology of Mandarin learners of English

Brook)	and the Gradual Learning Algorithm
Laura Colantoni & Jorge Gurlekian (University of Toronto and Laboratorio de Investigaciones sensoriales)	Early peak alignment and deep falls in Buenos Aires broad focus declaratives
Pauline Welby (Institut de la Communication Parlée, CNRS UMR 5009 Institut National Polytechnique de Grenoble, Université Stendhal Grenoble)	French Intonational Structure: Evidence from Tonal Alignment
Kyle E. Chambers, Kristine H. Onishi, Cynthia Fisher & Jessica Maye (University of Illinois and University of Rochester)	Learning different phonetic distributions in onset and coda position
Jose R. Benki (University of Michigan)	An investigation of front-back confusions in vowels
Stefan Benus, Adamantios Gafos & Louis Goldstein (New York University, Yale University and Haskins Laboratories)	Nonlinear links between continuity and discreteness: transparency in vowel harmony
Juli Cebrian (Universitat Autonoma de Barcelona)	Vowel phonotactics and syllabification in native and non-native phonology
Andries W. Coetzee (University of Massachusetts and Potchefstroom University)	Markedness from the grammar, not the lexicon
Susana Cortes, Mitsuhiko Ota & Alice Turk (University of Edinburgh)	Production of English Intervocalic /d/ by Catalan Speakers in Derived and Non-derived Words
Lisa Davidson (New York University)	Fitting phonotactics into gestural phonology: Evidence from non-native cluster production
Jeong-Im Han (Konkuk University)	Variability in Phonetic Implementation: Evidence from Korean Speakers' Perception of the Stop Epenthesis in English
Satomi Imai, James E. Flege & Ratree Wayland (University of Alabama at Birmingham and University of Florida)	Perception of cross-language vowel differences: A longitudinal study of native Spanish adults learning English
Mafuyu Kitahara, Tadao Murata & Haruka Fukazawa (Yamaguchi University and Kyushu Institute of Technology)	Articulatory correlates of sonority scale in consonants
Barbara Kühnert & Cécile Fougeron (Institut du Monde Anglophone and CNRS/UMR 7018 Université de la Sorbonne Nouvelle)	Schwa-productions and hesitations by native French speakers during English L2 performance
Shu-chen Ou & Mitsuhiko Ota (University of Edinburgh)	Metrical Computation in L2 Stress Acquisition: Evidence from Chinese-English Interlanguage
Ninik H. Poedjianto (University of Glasgow)	The Production of Indonesian English: Does L1 Phonological Transfer Occur At All?
Amanda Seidl & Eugene Buckley (Purdue University and University of Pennsylvania)	Infants' learning of unnatural rules
Minjung Son, Alexei Kochetov & Marianne Pouplier (Yale University, Haskins Laboratories, Simon Fraser University and University of Maryland)	The Role of Gestural Overlap in Perceptual Place Assimilation: Evidence from Korean
Geoffrey S. Nathan (Wayne State University)	Testing Usage-based Phonology, Just for the /l/ of it
Rajka Smiljanic & Ann Bradlow (Northwestern University)	Contrast enhancement in Croatian clear speech
Catherine Mayo & Alice Turk (University of Edinburgh)	The development of perceptual cue weighting within and across monosyllabic words
Tae-Jin Yoon, Heejin Kim & Sandra Chavarría (University of Illinois at Urbana-Champaign)	Local acoustic cues distinguish two levels of prosodic phrasing: Speech corpus evidence
Maksym Vakulenko (Institute for Theoretical Physics-Ukrajina and The University of Texas-Arlington)	Acoustic invariant approach and speaker-independent features of speech sounds (Ukrainian and English)

Steve Winters (Indiana University)	Why Some Things are Um-bearable: Investigating Perceptual and Articulatory Influences on Place Assimilation Processes
Alan C. L. Yu (University of Chicago)	Subphonemic tonal distinction in Cantonese
Elinor Payne (University of Cambridge)	Phonetic motifs and their role in the evolution of sound structure
Hyunsook Kang & Seok-Keun Kang (Hanyang University and Wonkwang University)	Processing English [s] into Korean: Prosodic influence on sound change
Chakir Zeroual (Centres des Etudes Universitaires de Taza and UMR 7018 CNRS/Sorbonne-Nouvelle)	A typological study of posterior articulations

University of Illinois at Urbana-Champaign Department of Linguistics Beckman Institute for Advanced Science & Technology

Exploring the Basis for Generalization in Language Acquisition

LouAnn Gerken* & Tania Zamuner†
University of Arizona and University of Nijmegen*†

Parameter Setting and other tightly constrained theories of language acquisition assume that the possible bases of generalization of all languages (Universal Grammar or UG) can be identified by studying the range of possibilities that exist across languages for a particular linguistic dimension (e.g., syllable shapes). Although this is an elegant theory, a variety of data have begun to suggest that it needs rethinking. I will discuss two types of data.

I'll begin by presenting toddler production data collected by Tania Zamuner, demonstrating that coda omissions do not appear to be governed by UG principles discoverable from cross-language differences. Rather, omissions are consistent with the view that children begin producing those codas that are phonotactically most typical of the language they are learning. These data cast doubt on one source of evidence for a link between cross-language differences and language acquisition by children.

I will then discuss a series of studies in which 9-month-olds are familiarized with a set of words illustrating four Optimality Theory (OT) stress assignment principles. A key property of the words is that they allow for a test of the transitivity principle of OT. Thus, we can ask if, presented with evidence that Principle A outranks Principle B and B outranks C, infants can infer that Principle A outranks Principle C. At test, infants are presented with new words that are either consistent or inconsistent with the A>>C ranking from their familiarization language. Two studies demonstrate that infants indeed discriminate the test stimuli, providing the first evidence that infants in the lab can generalize based on information extracted across multiple utterance types. However, an additional study calls into question whether the basis of infants' generalization is the one suggested by UG. Rather, infants, like the toddlers in the first study, appear to base their generalization on a typical property of the input stimuli. I will end by discussing new studies that explore the conditions under which generalizations of the type assumed by UG might be found.

The Prosodic Pattern of Words and Phrases in the Acquisition of German

Angela Grimm
University of Groningen

At the onset of word production, children are often claimed to show a trochaic bias (Allen & Hawkings 1980, Fikkert 1994, but see Demuth 1996, Vihman, de Paolis, & Davis 1998). Nevertheless, it is surely true for languages like German where the predominant prosodic pattern of words is a trochee, i.e. a disyllabic foot with initial stress (SW). In fact, the few acquisition data published so far and my own database of 5 children acquiring German indicate that German children prefer a SW template at the onset of meaningful speech.

In my data, the preponderance of trochaic feet persists beyond the first word combinations (= phrases), which emerge at around 18 months. In contrast to the single words, early phrases can contain more than two syllables and they show final prominence. Again, this pattern reflects the adult input where phrases get a nuclear accent in the unmarked case.

According to current models of prosodic development, the children would operate at the foot level as long as their output does not exceed the size of a foot (Fikkert 1994, Demuth & Fee 1995, Archibald 1995). The phonological structure of the early phrases, in contrast, shows that the children are *able* to produce different prosodic units at a time when single words let reason a stage with single feet only.

One implication is that children should have access to a phonological phrase level (where "phonological phrase" refers to any kind of prosodic level above the prosodic word) prior to the expansion of the word template. With the emergence of phrase level prosody at around 18 months, single words might now be bounded into phonological phrases. Phonetically, phrase boundaries are indicated by a declination of pitch or by phrase-final lengthening (PFL).

To ascertain whether single words and phrases do indeed provide evidence for phrase level constituents, two comparisons of the final rhymes have been performed on age-matched words and phrases with respect to the occurrence of (phrase final) lengthening. The results are can be summarized as follows: First, PFL appeared no matter whether the test items were single words (e.g. mam**a**) or phrases (da mam**a**). There was no significant difference in the duration of the word-final or phrase-final rhyme ($p > .05$, effect size .032). Second, the target words at phrase-initial position (pap**a** da) did not undergo PFL in contrast to single words (pap**a**), resulting in a significantly longer duration of the rhymes in single words ($p < .00$, effect size .051).

We conclude that the prosodic shape of a given word alone is not a sufficient indicator for the prosodic competence in child language. The prosodic form of the phrases as well as the phonetic evidence reveal that children do operate at a phonological phrase level earlier than a restriction on words only does suggest.

The present findings have consequences for current models of prosodic development: It can not longer be argued that level stress, arising by the collocation of two feet, is due to the absence of higher prosodic constituents. Moreover, since the prosodic hierarchy is already in evidence in phrasal constituents, level stress is not assumed to be part of the child grammar any more.

References

Allen, G.D. & Hawkins, S. (1980). Phonological Rhythm: Definition and Development. In. G. Yeni-Komshian, J. Kavanagh, & C.A. Ferguson, (Eds.), Child Phonology: Vol.1. Production, 227-256. New York: Academic Press.

Archibald, J. (1995). The Acquisition of Stress. In J. Archibald (Ed.) Phonological Acquisition and Phonological Theory, 81-109. Hillsdale, NJ: Lawrence Earlbaum Associates.

Demuth, K. & Fee, J. (1995). Minimal Words in Early Phonological Development. Ms., Brown University and Dalhousie University.

Demuth, K. (1996). The Prosodic Structure of Early words. In Morgan, J. & Demuth, K. (Eds.) Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition. 171-184. Lawrence Earlbaum.

Fikkert, P.M. (1994). On the Acquisition of Prosodic Sructure. Dordrecht, Holland Institute of Generative Linguistics.

Vihman, M. de Paolis, R. & Davis, B. (1998). Is there a "Trochaic Bias" in Early Word Learning? Evidence from Infant Production in English and French. Child Development 69 (4), 935-949.

Vocabulary growth and developmental expansion of types of phonological knowledge

Mary E. Beckman*, Benjamin Munson† & Jan Edwards*
Ohio State University and University of Minnesota†*

Studies of phonological processing over the last decade strongly support a model of adult mental lexicons in which the phonological form of each word is encoded in at least two ways. Each familiar wordform is encoded in terms of episodic memory traces of very fine-grained parametric representations of the auditory and articulatory patterns that are experienced in perceiving or producing specific instances of the word. This instance-based parametric encoding can account for indexical effects, such as the effects of hearing words produced by different voices on recall accuracy and reproduction at a delay (see, e.g., Goldinger, 1996). Positing such an instance-based parametric encoding also can account for token frequency effects, such as those described by Bybee (2000) and Jurafsky, Bell, and Girand (2002). The literature on type frequency effects, on the other hand, suggests a different encoding of each word's form, in terms of coarser-grained generalizations about phonological patterns that recur across instances of different words. For example, a large and growing literature shows that sequences of phonemes attested in many words (high-probability diphones) are processed and produced differently from those attested in few or no words (low-probability diphones). For example, diphone probability has been found to affect repetition latency for both real words and nonwords, but in an opposite direction (Vitevitch & Luce, 1999). Real words are named more slowly if they contain high-probability diphones, suggesting a competition effect from other similar words that have rich instance-based encodings at the parametric phonetic level. Nonwords, by contrast, are repeated more quickly if they contain high-probability diphones, suggesting that robust access to the fine-grained parametric representations relevant for producing a novel form is parasitic on the coarser-grained generalizations that are abstracted over encountering specific diphones in many different word forms.

If this model is correct, we should not be surprised to see evidence for some dissociation between the two types of encoding in phonological development. For example, we might expect to see deficits in the parametric phonetic encoding which are independent of any deficit in the coarser-grained generalizations. In this paper, we will review some relevant findings on phonological development in young children, focusing on several recent results from our laboratories. For example, we have found an effect of diphone probability on nonword repetition accuracy that is much more reliable in 3-7 year old children than in adults. Although the effect is smaller for older children, this interaction with age is completely subsumed by the size of vocabulary. Children with larger expressive vocabularies (which support more robust generalizations about production routines across words) show a much smaller effect of diphone probability. Perception accuracy in a gated word recognition task also is related to age in these typically-developing children, as well as to the presence of phonological disorders in another group of children tested at the same time. That is, among typically-developing children, younger children need more information than older children. Also, children with phonological disorders need more information than typically-developing age peers, suggesting a deficit in the parametric phonetic representations. Perception accuracy also predicts overall production accuracy, both in a word-naming task and in the nonword repetition task. However, neither perception accuracy nor production accuracy in the word-naming task predicts the size of the diphone probability effect once the size of the vocabulary is included in the regression model. These results predict

that children with specific language impairment, who have smaller vocabularies than their age peers with normal language development, should show a larger effect of diphone probability relative to their age peers with typical phonological development or to their age peers with phonological disorder without accompanying language delay. Preliminary results of another ongoing study suggest that this prediction will be confirmed.

References

Bybee, J. (2000). Lexicalization of sound change and alternating environments. Papers in Laboratory Phonology 5, 250-268.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. Journal of Experimental Psychology: Learning, Memory, & Cognition, 22, 1166-1183.

Jurafsky, D., Bell, A., & Girand, C. (2002). The role of the lemma in form variation. Laboratory Phonology 7, 3-34.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. Journal of Memory and Language, 40, 374-408.

The acquisition of abstract phoneme categories

Sharon Peperkamp*‡ & Emmanuel Dupoux*
Laboratoire de Sciences Cognitives et Psycholinguistique, EHESS - ENS – CNRS, Paris and Université de Paris 8‡*

There is much evidence that during the first year of life, infants acquire the segmental categories of their language. In a language's phonological system, not all segments have the same status. Specifically, different segments can be realizations either of different phonemes or of the same phoneme. In the former case, the contrast between them is phonemic, in the latter case it is allophonic. How could infants create abstract phoneme categories on the basis of concrete segmental categories? One possibility is that they use word meanings. For instance, once they know a number of different words that start with [b] and [d], they could infer that the contrast between [b] and [d] is phonemic in English. The presence of minimal pairs is particularly informative: knowledge that *ball* and *doll* are two different words suffices to conclude that there is a phonemic contrast between [b] and [d]. This learning algorithm thus requires the presence of a reasonably large lexicon. Alternatively, Peperkamp & Dupoux (2002) proposed that infants might construct phoneme categories prelexically on the basis of distributional information, since segments that are realizations of the same phoneme have complementary distributions. Thus, in Greek, [z] only occurs before voiced consonants, while [s] occurs elsewhere; this is evidence that [z] and [s] are realizations of a single phoneme. By contrast, [s] and [f] both occur, say, before vowels, hence they form a phonemic contrast.

In this talk, we consider the respective contributions of word meaning and distributional information for the acquisition of phoneme categories. In a series of artificial language-learning experiments, we show that French adults can learn to encompass voiced and voiceless stops or fricatives in abstract phoneme categories, whereas in their native language voicing is phonemic in all obstruents. We created two languages. In Language 1, voicing is phonemic in stops and allophonic in fricatives, with fricatives being voiced intervocalically. In Language 2, conversely, voicing is phonemic in fricatives and allophonic in stops. During exposure, subjects listen to short phrases in one of the two languages. According to the language, words starting with either a fricative or a stop occur in two phonetic forms, due to the intervocalic voicing rule. For instance, in Language 1, underlying /fami/ is realized as [vami] following vowel-final words and as [fami] elsewhere, while /puna/ is always realized as [puna], never as [buna] (which is in fact the realization of a different lexical item, /buna/); the reverse is true for Language 2. After exposure, subjects learn new words and are tested on their lexical representations of these words in a picture selection paradigm.

We ran three experiments that differed only in the exposure phase: In Exp. 1, all phrases were accompanied by a picture that allowed subjects to infer the meaning of the obstruent-initial words, and, according to the language, words with either initial stops or fricatives occurred in minimal pairs with respect to voicing of the initial consonant. Exp. 2 was identical to Exp. 1, except that there were no minimal pairs. In Exp. 3, subjects listened to the phrases in the absence of referential pictures; as in Exp. 2, there were no minimal pairs. In all three experiments, subjects treated word-initial stops and fricatives differently, in that they ignored voicing more often on the allophonic class than on the phonemic one for purposes of word identification. Whereas in the third experiment this difference was smaller than in the first two, there was no difference between subjects' performances in the first two experiments. These results suggest

that based on information regarding the distribution of segments, learners can create abstract phoneme categories that encompass different segments; lexical knowledge may be helpful but is not required, and the presence of minimal pairs is irrelevant.

We also report on a similar experiment that explores the role of natural classes during acquisition. In this experiment we test whether subjects generalize the acquired knowledge concerning phonemic versus allophonic voicing to obstruents that never occur intervocalically during exposure. We interpret and discuss our findings in light of current statistical learning approaches to first language acquisition.

Resyllabification revisited: the case of enchaînement in French

Cécile Fougeron
Laboratoire de Phonétique et Phonologie, UMR7018,CNRS/Univ. Paris 3, France

Over the last twenty years (or so), considerable evidence has been provided to support the claim that the prosodic organization of utterances conditions the phonetic surface form of a word. With the general objective of understanding the production mechanisms and the perceptual consequences of word form variations resulting from phrasal organization, this paper addresses the question of the phonetic basis of the so-called 'resyllabification' process across word boundaries.

There is a long tradition in the phonetic and phonological description of French to assume that the word does not have a phonetic reality and that word boundaries are blurred in the syllabification of the speech chain. Indeed, several external sandhi processes in French, like enchaînement and liaison, are said to result in the resyllabification of a word-final coda consonant with the following word-initial vowel in connected speech. Whatever the rationale for this post-lexical resyllabification -universal tendency to avoid onsetless syllables, general sonority principles, or language-specific rhythmic preferences- its consequence is the apparent creation of a recurrent string of CV syllables. The questions that are raised here are whether VC#V sequences are in fact phonetically resyllabified to a V.C#V structure, and whether cues of their abstract lexical and syllabic structure can be identified.

Several arguments based on acoustic analyses, perceptual results, as well as corpus-based statistics on cases of enchaînement will be provided to reject the view that the process of enchaînement implies a simple resyllabification of a word-final coda consonant to the onset of the following syllable, and induces a complete neutralization with sequences containing an underlying CV syllabic structure, regardless whether this CV syllable is word-initial or word-medial. Implications concerning the role of the syllable in the segmentation process in speech perception as well as in the phonological encoding in speech production will be discussed in light of these results.

Are tones aligned with articulatory events? Evidence from Italian and French

Mariapaola D'Imperio*, Hèléne Loevenbruck†, Caroline Menezes*, Noël Nguyen* & Pauline Welby†

Laboratoire Parole et Langage, Aix-en-Provence and Institut de la Communication Parlée, CNRS, Grenoble†*

Tonal alignment work has suggested that the temporal location of tonal targets relative to segmental "anchors" might be governed by principles of synchrony and stability (Arvaniti et al 1998, Ladd *et al.* 1999, *inter alia*). However, a number of discrepancies have emerged in the cross-linguistic study of alignment. For instance, despite some regularities in the alignment of L targets (Caspers and van Heuven 1993; Prieto *et al.* 1995), the alignment of H targets appears to be quite controversial. In fact, it is sometimes difficult to find definite segmental landmarks to which such targets might be aligned. Also, most of the alignment proposals so far inherently assume that if some anchors for tonal alignment do exist they must be acoustic in nature. A plausible alternative would be to assume that such anchors are primarily articulatory, which would explain why in some cases the underlying regularities would be masked. Hence, we adopt a new experimental paradigm for alignment research in which articulatory measures are performed simultaneously with acoustic measures.

In order to test the constant alignment hypothesis, a preliminary study (D'Imperio *et al.* 2003) was conducted in which various latency measures, both acoustically and articulatorily based, were analyzed. Specifically, the kinematics of OPTOTRAK markers attached to the speaker's upper and lower lip was tracked over time during the production of the corpus sentences. The melodic target considered is the H tone of LH nuclear rises in Neapolitan Italian. In this variety, yes/no question LH rises are systematically later than (narrow focus) statement LH rises (D'Imperio 2000, 2001, 2002; D'Imperio and House 1997). In order to test the hypothesis of constant anchoring of H targets, the materials were produced with two different rates of speech, i.e. normal and fast.

Summarizing the results, H targets of nuclear rises in Neapolitan statements and questions appear to be more closely phased with the articulatory dimension of between-lip distance than with two of the most commonly employed acoustic segmental landmarks for tonal alignment (i.e., onset and offset of stressed vowel). Statement H tones are phased with maximum between-lip distance within the stressed syllable. Note that this location does not correspond to any identifiable segmental boundary, acoustic event or phonological unit, and does not overlap with RMS peak amplitude. In fact, RMS peaks were generally much earlier than articulatory peaks, hence further away from H peaks. This calls for the collection and analysis of more articulatory data (especially jaw and tongue movements) to shed light on tonal alignment issues.

In a second study, a French corpus was collected on the basis of the alignment contrast found by Welby (2003, in press). Welby's results show that listeners use the alignment of the initial rise (LHi) in French Accentual Phrases as a cue to speech segmentation. Specifically, listeners exploit the presence of an early rise to demarcate the beginning of a content word. In the present study, a corpus was built with a set of utterances displaying this specific alignment contrast. The kinematics of 10 pellets (8 on the face and tongue, 2 references) was tracked over time using an electromagnetometer (EMA, Carstens). The phasing of several articulatory events relative to the L and H part of the early rise were examined.

The preliminary results seem to point to some kind of fine alignment specification for the L and H target. Specifically, we hypothesize that tonal target commands of Neapolitan as well as French rises are phased with commands of the supralaryngeal articulator involved to produce the segments to which the tone is associated. Regarding the word segmentation issue for French, it is important to study alignment in a diachronic perspective since we know of case of speech segmentation errors that can lead to lexical reinterpretation and change (*l'abondance* "abundance" > *la bondance*, from Welby 2003). We also take these results to suggest that not all rises align in the same way with the associated syllable. Though the role of articulatory constraints is important, the exact phasing properties of prosodic events are language-specific. Since prosody has recently become the realm of investigation of the Task Dynamics program (Byrd and Saltzman 2003), our alignment work will be cast under such a perspective.

Dialectal variation in the rising accents of American English

Gina Garding & Amalia Arvaniti
University of California, San Diego

In Pierrehumbert (1980) three rising accents are posited for English: H*, L+H* and L*+H. The need for three distinct accents has been disputed, however, most notably by Ladd (1983) who argued that they are all realizations of the same underlying accent under differing degrees of emphasis. Although all three accents have been investigated experimentally, such studies have tended to focus either on the scaling of the H under emphasis (e.g. Pierrehumbert, 1980; Ladd & Morton, 1997) or on the alignment of the tones in non-emphatic renditions (e.g. Pierrehumbert & Steele, 1989; Ladd & Schepman, 2003). Furthermore, in all of these experiments the dialect of the speakers was not controlled for.

The present experiment aimed to address these gaps. Our speakers belonged to two dialects, Midwestern and Southern Californian English. They read one-word utterances (consisting of the name 'Raymona') as part of two dialogues, one of which was designed to elicit L+H* accents and the other L*+H accents (cf. Hirschberg & Ward, 1992). The dialogues and our instructions prompted the speakers to use four levels of emphasis. If Pierrehumbert's analysis is correct, then the following results would be expected: (i) L+H* should show earlier alignment of both tones; (ii) the L tone of L*+H should get lower with added emphasis, while the L tone of L+H* should get higher under the same conditions. On the other hand, if Ladd (1983) is right, then non-emphatic renditions of our L+H* context should be realized as plain H* accents. Finally, experience with the two dialects at hand lead us to expect that any differences between the accents would be more pronounced in Midwestern than in Southern Californian English.

Our results so far confirm the general observation that peaks align later under emphasis (e.g. O'Connor & Arnold, 1973; Ladd & Morton, 1997). In all other respects, however, our results show two distinct patterns for the two dialects under investigation. Southern Californian speakers used H* in the non-emphatic tokens and L+H* in the emphatic tokens of both dialogues. Midwestern speakers, on the other hand, clearly differentiated L*+H from L+H*, and in both contexts showed evidence of a L tone at all levels of emphasis. In the Midwestern data, L*+H showed later alignment of both the L and the H tones. In addition, although the scaling of the H increased with emphasis in both accents, it was consistently lower in L*+H than in L+H*. The scaling of the L, on the other hand, was raised under emphasis only in L+H*; in contrast, in L*+H the L tone was consistently lower than in L+H*and remained relatively stable. Further, Midwestern speakers significantly lengthened the accented syllable when using L*+H, an effect not observed in the Southern Californian data.

These results suggest that Pierrehumbert's distinction between L*+H and L+H* is essentially correct, though it clearly does not apply to American English in general but only to some of its dialects. In addition, our results suggest that the distinction between H* and the L+H accents should be re-evaluated *for some dialects*, possibly along the lines advocated by Ladd & Schepman (2003) (and contra Ladd 1983). More generally, the behavior of L tones under emphasis corroborated Pierrehumbert's (1980) intuition (also supported in part by the perceptual evidence of Gussenhoven & Rietveld, 1997) that leading Ls are raised by subsequent Hs, but starred Ls are not. In this respect, our results show that with the right type of control, emphasis (which is often seen as paralinguistic and therefore outside the scope of linguistic description) can be a useful tool in the study of intonational phonology. Finally, the dialectal variation

uncovered by our study is strong evidence to the case that in order to fully understand the tonal makeup and role of accents in an intonational system, it is essential that experimental work does not overlook the dialect variable.

References

Gussenhoven, C. & T. Rietveld (1997) Empirical evidence for the contrast between H* and L* in Dutch rising contours. In A. Botinis, G. Kouroupetroglou & G. Carayiannis (eds.), Intonation: Theory, models and applications: Proceedings of an ESCA Workshop. Athens, Grenoble: ESCA.

Hirschberg, J. & G. Ward (1992) The influence of pitch range, duration, amplitude and spectral features on the interpretation of the rise-fall-rise intonation contour in English. Journal of Phonetics 20: 241-251.

Ladd, D. R. (1983) Phonological features of intonational peaks. Language 59: 721-759.

Ladd, D. R. & R. Morton (1997) The perception of intonation emphasis: continuous or categorical? Journal of Phonetics 25: 313-342.

Ladd, D. R. & A. Schepman (2003) "Sagging transitions" between high pitch accents in English: experimental evidence. Journal of Phonetics 31: 81-112.

O'Connor, J. D. & G. F. Arnold (1973) Intonation of Colloquial English. London: Longman.

Pierrehumbert, J. (1980) The Phonology and Phonetics of English Intonation. MIT Ph.D. dissertation.

Pierrehumbert, J. B. & S. A. Steele (1989) Categories of tonal alignment in English. Phonetica 46: 181-196.

Fundamental frequency extrema on weak syllables affect the relative prominence of strong syllables

Laura Dilley
MIT and Harvard University

In standard autosegmental-metrical (AM) analyses of English intonation (cf. Pierrehumbert 1980) two distinct types of analyses are offered for fundamental frequency (F0) maxima and minima (i.e., extrema) occurring on weak syllables. Phonological analyses treat such points as direct reflexes of underlying phonological elements, as when an F0 maximum on a weak syllable preceding a strong syllable is attributed to an unstarred H tone of a H+!H* pitch accent (Beckman and Ayers-Elam 1993). Phonetic treatments are also possible, in which F0 extrema are analyzed as the consequence of processes of phonetic interpolation or as within-category phonetic variation (Ladd and Schepman 2003, Silverman and Pierrehumbert 1990). These two types of analyses are sometimes in conflict (Ladd 2000). In a series of experiments testing predictions of the standard AM model, Dilley (forthcoming) probes the status of F0 maxima and minima as phonological vs. phonetic. Stimuli for these experiments were designed by shifting an F0 maximum or minimum in equal temporal increments through syllable sequences with specific stress patterns. Earlier production data obtained by Dilley in an imitation task (reported as Redi 2003) showed categorical effects, thus supporting a phonological analysis for F0 extrema. While designing stimuli for follow-up experiments, it was noticed that the alignment of an F0 maximum or minimum seemed to influence the relative prominence in words with more than one possible location for main stress, e.g. lémonade vs. lemonáde. Four stimulus series were created by (1) shifting an F0 maximum in equal increments through the SW sequence in millionaire and the SWW sequence in Lannameraine; and (2) shifting an F0 minimum in equal increments through the SW sequence in lemonade and the WS sequence in nonrenewable. All stimuli were created using Praat software (Boersma and Weenik 2000), and each target word was embedded in a short carrier phrase. Three experiments were conducted using these stimuli. In two complementary experiments utilizing discrimination and imitation tasks, the results reported by Redi (2003) were replicated and extended. In addition, a third experiment utilized stimulus continua based on the words millionaire, Lannameraine, and lemonade to determine how shifting an F0 extremum through a SW or SWW syllable sequence affected the perceived relative prominence of stressed syllables in these words. In a forced-choice task, 19 listeners judged whether the first or last stressed syllable in each stimulus item was the strongest. Results showed a quasi-categorical shift along each stimulus continuum in the rate of reporting the last syllable as the strongest. When the extremum was aligned with the initial stressed syllable of any of the words, listeners reported the first syllable to be the strongest. When the extremum was aligned with (1) the weak syllable in millionaire or lemonade, or (2) the second weak syllable in Lannameraine, listeners reported that the last syllable was the strongest. When the maximum was in the first weak syllable of Lannameraine, the majority of listeners reported that the first syllable was the strongest. These results support the hypothesis that F0 extrema should be analyzed as phonological, since their alignment influences the perceived location of main stress. Finally, casual observation suggests that, for words in which the position of the primary and secondary stresses are fixed, such as mármalàde, alignment of an F0 extremum with the weak syllable may be ill-formed. A follow-up study is planned to test this possibility. We hypothesize that the alignment of an F0 extremum with a weak syllable preceding a syllable with fixed secondary

stress represents an unstable configuration which is likely to lead to diachronic change in the stress pattern(s) of lexical items.

References

Beckman, M. and Ayers-Elam, G. (1997) "Guidelines for ToBI labeling," www.ling.ohio-state.edu/research/ phonetics/E_ToBI.

Boersma, P., and Weenik, D. (2000) Praat, a system for doing phonetics by computer, Version 3.4. Technical Report 132, Institute of Phonetic Sciences of the University of Amsterdam. www.praat.org.

Dilley, L. (forthcoming) Ph.D. Thesis, MIT and Harvard University.

Ladd, D.R. (2000) Tones and turning points: Bruce, Pierrehumbert, and the elements of intonational phonology. In M. Horne (ed.), Prosody: Theory and Experiment - Studies presented to Gösta Bruce, Dordrecht: Kluwer, pp. 37-50.

Ladd, D.R., and Schepman, A. (2003) "Sagging transitions" between high accent peaks in English: experimental evidence. Journal of Phonetics 31: 81-112.

Pierrehumbert, J. (1980) The phonology and phonetics of English intonation. Ph.D. Thesis, MIT.

Redi, L. (2003) Categorical effects in production of pitch contours in English. In Proceedings of the 15th International Congress of the Phonetic Sciences, Barcelona, p. 2921-2924.

Silverman, K. and Pierrehumbert, J. (1990) The timing of prenuclear high accents in English. In Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech. J. Kingston and M. Beckman (eds.), Cambridge University Press, Cambridge, 71-106.

Posterior Tongue Body constriction location differences in Click types

Amanda Miller-Ockhuizen*, Levi Namaseb*†T & Khalil Iskarous‡
Cornell University, University of Namibia†, University of TorontoT and Haskins Laboratory‡*

In this study, we investigate the articulation of the palatal [=] and post-alveolar [!] click types in Khoekhoe through simultaneous ultrasound imaging and dynamic electro-palatography. Click production involves two tongue constriction gestures, one using the tongue tip or blade, and another using the tongue body, which in combination result in the creation of a cavity in the oral tract, used to rarify air by creating suction through a velaric airstream mechanism. Click type labels generally derive from the anterior constriction location produced with the tongue tip for the post-alveolar click type [!], and the tongue blade for the palatal click type [=]. The anterior constrictions of clicks have been better documented than the posterior constrictions given the difficulty of capturing the back place of articulation in traditional as well as electro palatography. We focus on the investigation of the posterior constriction location in the palatal [=] and post-alveolar [!] click types at the moment just prior to release in the [i] and [ai] contexts, and normalize the individual productions relative to [i].

Cineradiology data published in Traill (1985:110) for a related language !Xoo show velar posterior tongue body constriction locations for all click types, and this forms the basis of the tracings outlined in Ladefoged and Traill (1984) and Ladefoged and Maddieson (1996: 250). However, based on the cross-linguistic patterning of [!] in Khoe and San languages with uvularized consonants and epiglottalized consonants in retracting and lowering [i] (Miller-Ockhuizen 2000, 2003), we hypothesize that the post-alveolar [!] click type contains a uvular constriction produced with tongue body retraction, while the palatal click [=] involves tongue root raising into the upper pharyngeal region. This hypothesis differs from the findings of Traill (1985) where the posterior constriction in all click types is velar. The presumed unified velar place of the posterior constriction in all click types has confounded our understanding of the phonology of clicks in a process of lowering and retraction of [i] called the Back Vowel Constraint (Traill 1985: Miller-Ockhuizen 2000, 2003), and left us with no phonetic basis for the generalization found by Miller-Ockhuizen (2003) that the post-alveolar [!] click type patterns together with uvulars.

Ultrasound images and EPG contact patterns confirm that the post-alveolar [!] click type has a uvular gesture produced with the tongue body to form the posterior constriction of this click type. The phonological incompatibility of the post-alveolar click type [!] with the front vowel is due to the raising and retraction of the back of the dorsum in the uvular region to form the posterior edge of the velaric cavity. The uvular gesture would use the styloglossus, which is an antagonist of genioglossus, needed for the palatal gesture in the production of the front vowel [i]. The phonological compatibility of the palatal [ʔ] click type with the front vowel, on the other hand, can be understood in terms of the dynamics of the tongue. In order to raise the tongue root to make a constriction in the upper pharyngeal region, hyoid raising muscles like geniohyoid are recruited. An outcome of geniohyoid contraction is the raising of the front of the tongue, which assists in the accomplishment of the palatal gesture for the front vowel [i]. The higher position of the center of the tongue in the palatal click type also accounts for the degree of fronting coarticulation found in the vowel [a] in Khoekhoe, and in Ju|'hoansi (Miller-Ockhuizen 2001), which was formerly referred to as Dental Assimilation by Sagey (1986).

Crucially, the tongue root retraction articulated with the uvular gesture in [!] provides an understanding of the cross-linguistic patterning of [!] with uvulars. In Khoekhoe, the incompatibility of [!] with [i] shows up in the low frequency of this sequence in the lexicon compared with the high frequency of sequences of [=] followed by [i]. The retraction and lowering of [i] associated with this click type and other pharyngeals can now be seen as a common cross-linguistic process of pharyngeal assimilation that is categorical in languages like !Xoo and Ju|'hoansi. Frequency based effects can become allophonic through gradual change in the lexicon. Moreover, the palatal tongue body constriction location in [=] correlates well with the higher F2 values seen following this click type in Khoekhoe (this study) and in Ju|'hoansi (Miller-Ockhuizen 2000). This process was originally termed Dental Assimilation by Sagey (1985), but was shown to be a more gradient coarticulatory phenomenon by Miller-Ockhuizen (2000). In this study, we show that it involves fronting of the tongue and is not restricted to the dental region.

References

Ladefoged, P. and Maddison, I. (1996) The Sounds of the World's Languages, Cambridge, Blackwell Publishers.

Ladefoged, P. and Traill, A. (1984) Linguistic Phonetic Descriptions of Clicks, Language 60 (1), 1-20.

Miller-Ockhuizen, A. (2000) C -V Coarticulation and Complex Consonants: Evidence for Ordering in Click Place Gestures, In Proceedings of LP '98: Item Order in Language and Speech (Fujimura, O., Joseph, B. & Palek, B., editors), Prague, Charles University Press.

Miller-Ockhuizen, A. (2003) The Phonetics and Phonology of Gutturals: A Case Study from Ju|'hoansi, Outstanding Dissertations in Linguistics, Routledge.

Sagey, Elizabeth (1986) Representation of Features and Relations in Non-Linear Phonology, MIT Dissertation in Linguistics.

Traill, A. (1985) Phonetic and Phonological Studies of ✍X???Bushman (Quellen zur Khoisan-Forschung 1), Hamburg, Helmut Buske Verlag.

Phonetic reflexes of morphological boundaries under different speech rates

Mariko Sugahara & Alice Turk
University of Edinburgh

One of the universal aspects of speech is that speakers manipulate acoustic parameters to mark linguistic boundaries. For example, it is well-documented that word-final syllables are longer when immediately followed by sentence/utterance and phrase boundaries than when followed by no such boundaries (Cooper & Paccia-Cooper 1980; Whiteman, Sahttuck-Huffnagel, Ostendorf & Price 1992; among many others). Recent work by Turk and colleagues (Turk & Sawusch 1997; Turk & White 1999; Turk & Shattuck-Hufnagel 2000) and Beckman & Edwards (1990) have further shown that acoustic duration of segments and syllables is also affected by the presence or absence of a word-level constituent boundaries (i.e. pre-word boundary syllables are longer than word-medial syllables). However, reading materials used in those near-word boundary duration adjustment studies were limited to a set of two target words forming a morphological compound word or a phrase. Therefore, it is not yet clear whether the "near-word boundary duration adjustment" is confined to word sequences or more generally to any forms with a morphological boundary. One of the main goals of this paper is to examine the pre sence or absence of boundary duration adjustments in forms containing Level II suffixes in English. Also we will examine how different speech rates affect the duration adjustment near Level II suffix boundaries. It is already known that a near-word boundary durational adjustment depends on various contexts such as the presence or absence of a pitch accent (Turk & S-H) and a speech rate (Beckman & Edwards). Therefore, different speech rates may result in different quantity/quality of durational adjustments near a Level II suffix boundary.

In English, affixes belong to two classes, the less productive Level I affix class (*-al, -ion, -ic*, etc) and the more productive Level II affix class (*-ing, -s, -er, -ness*, etc). Siegel (1974), Selkirk (1982), Borowsky (1993) and Aronoff & Srider (1983) agree based on phonological evidence that a boundary between the stem and the Level II suffix is stronger than that between the stem and the Level I suffix. It is worthwhile testing whether the phonologically motivated "stronger" Level II suffix boundary has any phonetic reality. In our first set of experiments using a normal speech rate only, we compared the duration of (a) a rhyme immediately followed by the Level II suffix boundary (i.e. a singular person/a possessive suffix *–s*, a past tense suffix *–t*, an agentive suffix *–er*, and a progressive suffix *-ing*) and (b) that of a corresponding rhyme within a monomorphemic word and (c) that of a rhyme immediately followed by a word boundary. We found slightly but significantly longer rhyme duration in a subset of the Level II suffix boundary cases (i.e. in the suffixed forms containing *–s* or *–t* but not in those containing *–er* or *–ing*) than in the monomorphemic cases. This result was consistent with the results obtained in Schwarzlose & Bradlow (2001), who compared the duration of a segment immediately preceded by a plural and a first person singular *–s* and that of a corresponding segment in a monomorphemic context. Though it was slightly longer than the monomorphemic case, the duration of the rhyme of the suffix cases was still shorter than that of the word boundary cases. In our next set of experiments, we will compare the duration of forms containing those four suffixes (i.e. *–s, -t, -er, -ing*) and that of their monomorphemic counterparts read in a slower speech rate, which we are preparing for the conference. We assume following Beckman & Edwards' observation that a slower speech enhances duration markers of a morphological/prosodic boundary. If the subtle durational difference found in the first

experiment is a reflex of a morphological/prosodic boundary available in a phonetic component, we expect a more consistent and larger difference to emerge under the slower speech rate condition.

References

Aronoff, M. & Sridhar, S. N., 1983. Morphological Leels in English and Kannada; or Atarizing Reagan. Papers from the Parasession on the Interplay of Phonology, Morphology and Syntax, Chicago Linguistic Society, 3-16.

Beckman, M. E. & Edwards, J., 1990. Lengthening and Shortening of Prosodic Constituency. In Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech, J. Kingston & M. E. Backman (eds.). Cambridge: Cambridge University Press.

Borowsky, T., 1993. On the Word Level, In Phonetics & Phonology: Studies in Lexical Phonology, S Hargus & E. M. Kaisse (eds.). Academic Press.

Cooper, W. & Paccia-Cooper, J., Syntax and Speech. Cambridge, MA: Harvard University Press.

Schwarzlose, R. & Bradlow, A. 2001. What happens to segement durations at the end of a word? Journal of the Acoustical Society of America, 109(5, pt.2), 2292.

Selkirk, E. 1982, The Syntax of Words. Cambridge, MA: MIT Press.

Siegel, D. 1974, Topics in English Morphology. Doctoral dissertation, MIT.

Turk, A. & Shattuck-Hufnagel, S., 2000. Word-Boundary-Related Duration Patterns in English. Journal of Phonetics, 28. 397-440.

Turk, A. & Sawusch, J., 1997. The Domain of Accentual Lengthening in American English, Journal of Phonetics, 25, 25-41.

Turk, A. & White, L. 1999. Structural Influences on Accentual Lengthening in English, Journal of Phonetics, 27(2), 171-206

Whiteman, C. W., Shattuck-Hufnagel, S., Ostendorf, M. & Price, P.J, 1992. Segmental Durations in the Vicinity of Prosodic Phrase Boundaries, Journal of the Acoustical Society of America, 90(3), 1707-1717.

Speech in its natural environment: accounting for social factors in phonetic variability

G.J. Docherty
University of Newcastle upon Tyne

While much work under the rubric of Laboratory Phonology has been concerned with phonetic variability, how it is manifested in speech production and processed in perception, conspicuously little attention has been given to those aspects of inter- and intra-speaker performance variability which correlate systematically with a range of social factors characterising speech communities and individual speakers.

Nevertheless, recent theoretical models suggest that such social-indexical properties of speech may need to be accounted for within models phonological knowledge as much as those properties relating to lexical contrast. In pursuit of this objective, this paper presents an overview of social factors in speech production. Drawing on a range of recent studies, some key outstanding questions are identified and discussed, and pointers are given to the ways in which theoretical frameworks will have to evolve if they are to accommodate social-indexical aspects of speech.

Empirical evidence for gendered speech production: /s/ in Glaswegian

Jane Stuart-Smith
University of Glasgow

It is well established that the acoustic characteristics of /s/ differ in male and female speakers, with females typically showing higher frequency energy than males (e.g. Flipsen et al 1999). These differences are often accounted for by referring to the observation that female speakers are likely to have a shorter resonance cavity in front of the fricative constriction than males (e.g. Stevens 1998). Such explanations emphasize these differences in terms of 'sex', i.e. the biological determination of being male or female, as opposed to 'gender', i.e. the socio-cultural construct of being male or female. However, Strand (1999) notes male/female vocal tract size differences are largely behind the area of fricative constriction, and argues for 'socially influenced fricative productions'. This moves towards the suggestion that males and females may use different strategies to articulate /s/ as part of a more general construction of themselves as 'masculine' or 'feminine' within the culture and society in which they live. This in turn fits in with the recognition that social-indexical information is carried in the speech stream along with linguistic information (e.g. Docherty 2003; Abercrombie 1967); thus we should expect to find gender represented amongst the social-indexical features of /s/ alongside those which may be determined by sex. However, to date there appears to be no empirical evidence to substantiate the claim that female speakers are using specific strategies to produce /s/. To test this, it is necessary to have access to female (and male) speakers of the same language but who belong to different socio-cultural groups, and so who might be expected to show different norms for constructing gender, which may possibly extend to the production of /s/.

The accent of the city of Glasgow in Scotland offers just such a case. Glasgow is known for showing sharp socio-cultural differences between the middle- and working-classes, which is reflected in clear sociolinguistic stratification. Moreover, an apico-alveolar articulation of /s/ is associated with vernacular Glaswegian (e.g. Macafee 1983: 34), though preliminary auditory analysis of /s/ in a socially-stratified corpus of Glaswegian speech confirmed this largely for younger male speakers in general (Stuart-Smith, 1999: 209).

This paper presents the results of an acoustic analysis of Glaswegian /s/, and tackles three key research questions. 1) What are the acoustic characteristics of /s/ in Glaswegian? 2) Do male and female speakers differ according to 'sex', and in the expected direction, i.e. females showing higher frequencies than males? 3) Is there evidence for gender differences? The data for this study consist of digital recordings of read speech from 32 speakers, young and old, male and female, working-class and middle-class. All wordlists were digitised into a CSL running MultiSpeech at 48,000 Hz with 16 bits, with subsequent high-pass and low-pass filtering. /s/ was analysed in a selection of words which varied according to position of /s/ in the word and vowel context. Acoustic analysis was based on wide-band spectrograms and time-averaged spectra of at least 7 spectra per fricative (512-point DFT, Hamming window), taken along the duration of the fricative. A range of spectral measurements were taken. Following the work of Shadle (e.g. Jesus and Shadle 2002), I took the frequency of the highest amplitude peak, and the spectral slopes leading up to and falling away from the peak, since these are predicted to vary both according to place of articulation and, slightly differently, to a speaker's sex. Four measures from a moment analysis (cf Forrest et al 1988) were also used to capture the overall distribution of spectral energy.

The results show consistently higher frequency peaks/first moment for most, but not all, female speakers. Working-class girls, as a group pattern with male speakers, and there are other indications of gender in both male and female speakers. Moreover, the distribution of slope measures confirm this finding, and also indicate differences of age/sex in speakers. Overall, this study provides the first empirical evidence that aspects of the spectral energy of /s/ are modified by speakers to signal gender.

References

Abercrombie, D. (1967), Elements of General Phonetics, Edinburgh: EUP

Docherty, G. (2003), 'Speaker, community, identity: Empirical and theoretical perspectives on sociophonetic variation', Proceedings of 15th International Congress of Phonetic Sciences, Barcelona

Flipsen, P., Shriberg, L., Weismer, G., Karlsson, H. and McSweeny, J. (1999), 'Acoustic characteristics of /s/ in adolescents', Journal of Speech, Language, and Hearing Research, 42, 663-77

Forrest, K., Weismer, G., Milenkovic, P. and Dougall, R.N. (1988), 'Statistical analysis of word-initial voiceless obstruents' JASA, 84, 115-123

Jesus, L. and Shadle, C. (2002), 'A parametric study of the spectral characteristics of European Portuguese fricatives', Journal of Phonetics, 30, 437-64

Macafee, C. (1983), Glasgow, Amsterdam: Benjamins

Stevens, K. (1998), Acoustic Phonetics, Cambridge MS: MIT Press

Strand, E. (1999), 'Uncovering the role of gender stereotypes in speech perception', Journal of Language and Social Psychology, 18, 86-99

Stuart-Smith, J. (1999) 'Glasgow: Accent and voice quality' in P. Foulkes and G. Doherty (eds), Urban voices: Accent studies in the British Isles, Edward Arnold, 201-22

Origins of Canadian Raising in Voiceless-Coda Effects: A Case Study in Phonologization

Elliott Moreton* & Erik Thomas†
University of North Carolina at Chapel Hill and North Carolina State University†*

Phonological alternations are often taken to be "phonologized" versions of more subtle phonetic effects. This study traces the development of one well-known English phonological alternation, Canadian Raising ("CR") of /ai/, documenting the phonetic basis and the course of phonologization with instrumental data from speakers born as long ago as 1878. The findings contradict the widely accepted view that CR is a consequence of the Great Vowel Shift (e.g., Trudgill 1986:158–161, Donegan 1992, Britain 1997, Stockwell & Minkova 1997), provide an extremely detailed picture of the origin of a phonological alternation, and suggest that one phonological rule may embody the effects of more than one phonetic process.

CR is usually (incorrectly) characterized as the raising of diphthong nuclei before voiceless codas: [t^It] *tight* versus [taId, taI, taIm] *tide, tie, time.* In Canada /au/ is affected as well, but this study focuses on the more widespread raising of /ai/. CR is a nearly paradigmatic case of a phonological rule. It is regular and productive. It is categorical in that native speakers have clear judgments, and that there are lexical exceptions (Vance 1987).

We propose that CR is a phonologization of the phonetic result of three simultaneous processes. Before a voiceless coda, vowel articulations tend to be peripheralized in vowel space (Wolf 1978, Summers 1987, Thomas 2000; Moreton in press), while diphthong nuclei are abbreviated (Gay 1968). The peripheralized /ai/ offglide raises the weakened nucleus through coarticulation. The phonological rule could then have arisen when this was reinterpreted as a target difference ("hypocorrection", Ohala 1981, 1993). The following arguments are advanced:

(1) /ai/ glides are more peripheral before voiceless obstruents in every English dialect in which they have been examined instrumentally (Thomas 2000, 2001). This relationship also holds for /oi/ and /ei/ over a wide range of U.S. dialects (Moreton in press).

(2) Raising is in fact more widespread and more prominent in the offglide than in the nucleus. This is true not only for /ai/ in modern CR dialects, but also for /oi/ and /ei/ generally (Moreton in press 2003). The Great Vowel Shift, on the other hand, produced /ai/ by lowering the nucleus of Middle English /i:/, not the offglide, and did not affect /oi/.

(3) "Canadian" Raising has been repeatedly re-innovated around the English-speaking world (Trudgill 1986:160, Britain 1997). By measuring archival and recent recordings, we have documented one of those re-innovations in the Western Reserve area of Ohio. The Western Reserve is relatively insulated from Canadian influence: U.S. dialects resist a range of Canadian variants (Labov et al. forthcoming Ch. 11, Zeller 1993), documented innovations in the Western Reserve have not come from Canada (Drake 1961, Thomas 2001), and the major sources of migrants to the Western Reserve have not been CR areas. Speakers born in the late 19th Century have the universal pattern of no nuclear raising and slight offglide raising, which gradually progresses to raising of Canadian magnitude among those born in the late 20th Century.

(4) Many varieties of African-American English and Southern (white) U.S. English have an alternation in /ai/ which parallels CR: [taIt] *tight* versus [ta:d, ta:, ta:m] *tide, tie, time.* Paradoxically, the more-demanding diphthong articulation appears in the short pre-voiceless environment (Sledd 1966). It makes phonetic sense only in terms of the present theory: The long ("drawled") nuclei of these dialects are better able to resist the coarticulatory pull of the

offglide. The pre-voiceless offglides can thus hold their own, but the elsewhere offglides, unaided by peripheralization, assimilate to the nucleus.

This study extends a phonologization literature which has focused on splits and mergers rather than alternations, documents the emergence of an alternation in detail, and shows how one phonological rule can combine multiple phonetic effects.

References

Britain, David (1997) Dialect contact and phonological reallocation: "Canadian Raising" in the English Fens. Language in Society, 26, 15–46.

Donegan, Patricia (1993). On the phonetic basis of phonological change. In: Charles Jones (ed.), Historical Linguistics: Problems and Perspectives, pp. 98–130. London: Longman.

Drake, James A. (1961). The effect of urbanization on regional vocabulary. American Speech 36:17-33.

Gay, Thomas (1968). Effect of speaking rate on diphthong formant movements. Journal of the Acoustical Society of America 44:1570–1573.

Labov, William, Sharon Ash, and Charles Boberg (in preparation). Phonological Atlas of North American English. Berlin: Mouton de Gruyter. www.ling.upenn.edu/phono_atlas/home.html

Moreton, Elliott (in press). Realization of the English postvocalic [±voice] contrast in F1 and F2. To appear in Journal of Phonetics.

Ohala, John J. (1981) The listener as a source of sound change. In: Carrie S. Masek, Roberta A. Hendrick, and Mary Frances Miller (eds.), Papers from the Parasession on Language and Behavior, Chicago Linguistic Society, May 1-2, 1981, 178-203. Chicago: Chicago Linguistic Society.

Ohala, John J. (1993). The phonetics of sound change. In: Charles Jones (ed.), Historical Linguistics: Problems and Perspectives, 237–278. London: Longman.

Sledd, James H. (1966). Breaking, umlaut, and the Southern drawl. Language 42:18–41.

Stockwell, Robert, and Donka Minkova (1997). On drifts and shifts. Studia Anglica Posnaniensia 31:283-303.

Summers, W. Van (1987). Effects of stress and final-consonant voicing on vowel production: Articulatory and acoustic analysis. Journal of the Acoustical Society of America 82(3):847-863.

Thomas, Erik R. (2000). Spectral differences in /ai/ offsets conditioned by voicing of the following consonant. Journal of Phonetics 28(1):1-25.

Thomas, Erik R. (2001). An Acoustic Analysis of Vowel Variation in New World English. Publication of the American Dialect Society 85. Durham, NC: Duke University Press.

Trudgill, Peter (1986). Dialects in Contact. New York: Blackwell.

Vance, Timothy J. (1987). "Canadian Raising" in some dialects of the northern United States. American Speech 62:195-210.

Wolf, Catherine G. (1978). Voicing cues in English final stops. Journal of Phonetics 6:299–309.

Zeller, Christine. 1993. Linguistic symmetries, asymmetries, and border effects within a Canadian/American sample. In Sandra Clarke (ed.), Focus on Canada. Amsterdam/Philadelphia: John Benjamins.

The loci of sound change effects in recognition and perception

Paul Warren*, Jen Hay† & Brynmor Thomas†
Victoria University of Wellington and University of Canterbury†*

Most models of lexical representation have been based on the results of tightly controlled experiments. Thus, while it is clear that socially conditioned variation exists and is highly structured, most models of lexical representation and speech perception are not well-equipped to deal with this variation. Our research programme investigates the processes of spoken word recognition in the context of a system that is in a state of flux. The particular situation that we investigate is the merger in New Zealand English of the NEAR and SQUARE diphthongs, a merger of approximation on the NEAR vowel that has been proceeding through the variety over the last twenty years or so. We report on an initial series of speech production and perception tasks run in Christchurch and Wellington:

a) timed lexical decision tasks exploring the issue of differential access to NEAR and SQUARE words on exposure to NEAR or SQUARE tokens;
b) the outcomes of binary forced-choice identification between NEAR and SQUARE words;
c) variants of (b) which manipulate the perceived age of the speaker in order to investigate how disambiguation may be affected by a listener's knowledge of speaker characteristics;
d) acoustic analyses of word-list data from all participants.

In this paper, we bring together the results of these studies and consider their implications for models of spoken word recognition. These studies demonstrate:

(i) listeners' lexical access processes reflect their experience of both merged and unmerged systems in the environment;
(ii) these lexical access processes show sensitivity to perceived speaker attributes;
(iii) listeners who merge the vowels in their own speech, and report that they cannot hear the difference between them in the speech of others, are nevertheless able to discriminate between the vowels reliably in a forced-choice identification task;
(iv) their accuracy at this task is related to the extent to which there is still some separation in their own speech;
(v) participants' ability to identify members of word-pairs correlates not with how distinct the stimulus words are, but how distinct that word-pair tends to be kept in the real world. Given two word-pairs which are equally phonetically distinct, individuals are likely to perform more accurately on the word-pair which has been produced most distinctly in their accumulated experience;
(vi) listeners also show sensitivity in their perceptual discrimination to perceived speaker attributes, both those based on characteristics of the speaker's voice and those induced by age-differentiating photographs purportedly of the speaker;
(vii) SQUARE words are more prone to error in identification tasks than NEAR words.

This set of results can be best accounted for by experience-based models of speech perception, in which the representation of a particular word is a distribution of remembered exemplars, complete with phonetic detail (Johnson 1997, Pierrehumbert 2001). These stored exemplars bias lexical access, so listeners are better able to attend to the distinction in word-pairs that are more often kept distinct (i, iii, iv, v). The set of prior experiences must include implicit knowledge of how variation is socially distributed (ii, vi). However the set of results cannot be accounted for by an episodic lexicon alone (c.f. vii). A rapid phonological pre-processor is also

required. Because the merger in NZE is towards a close-starting point for both diphthongs, this form is more frequent in the participants' environment. Hence, a "fast phonological pre-processor" (Pierrehumbert 2001) would be biased towards the closer diphthong, so that SQUARE tokens may occasionally be heard as NEAR (vii). Taken together, this set of results provides strong evidence for an experience-based model of spoken word recognition, with both lexical and prelexical levels.

References

Johnson, K. (1997). Speech perception without speaker normalization. In J. W. Mullennix (Ed.), Talker Variability in Speech Processing (pp. 145-166). New York: Academic Press.

Pierrehumbert, J. (2001). Why phonological constraints are so coarse-grained. Language and Cognitive Processes, 16, 691-698.

Evidence for a relationship between synchronic variability and diachronic change in the Queen's annual Christmas broadcasts.

Jonathan Harrington
University of Kiel

There is considerable evidence that sound change is initiated by the young, but the mechanisms by which changes in progress propagate through the community are not well understood. In one of the models proposed by Labov (2001), shifts in pronunciation are incrementally propagated in waves through generations because of sound changes that take place in individuals below a certain critical age. However, a small number of real-time studies provide evidence that adult speakers are not immune from sound changes going on around them. This would suggest that the phonetic exponents of phonological categories are not hard-wired in adulthood but that they can instead be updated in the light of community changes. In previous research (Harrington, 2000a, b, in press), we investigated the stability of vowel change in adults by carrying out a real-time analysis of the vowels from the annual Christmas broadcasts of Queen Elizabeth II over three time intervals between the 1950s and the 1980s and have found that both monophthongs and diphthongs from the 1950s are shifted in the direction of a more modern and less aristrocratic form of Received Pronunciation. In the research to be reported here, we extend these results both by tracking the annual shift in selected monophthongs during the period in which change had been found to be greatest (roughly from the late 1950s to the early 1970s) and by analyses of broadcasts from recent years (1999-2002). With these data, we test the extent to which there is a relationship between sychronic variability and diachronic vowel shift by analysing two vowels, /u/ and /æ/ which are known to have changed in the community since the 1950s. The data spanning a fifty year period allows us to explore whether chain shifting, which has been extensively documented in both historical linguistics and more recently in sociolinguistic investigations of apparent time studies across different age groups, also underlies the diachronic change of vowel positions within the same person. The results of our analysis so far support usage-based and exemplar models of speech production (Bybee, 2001; Pierrehumbert, 2001) in which phonological categories in adulthood can be selectively and incrementally updated by adaptation to changes taking place in the community.

Aerodynamics of [r] in tonogenesis

Susan G. Guion* & Ratree P. Wayland†
University of Oregon and University of Florida at Gainesville†*

It has been noted by several scholars that dialects of Khmer (Cambodian) have developed a falling tonal contour in words containing an apical trilled [r] in the onset [1,2,3,4,5]. The goal of the proposed paper is to investigate a possible phonetic motivation for this tonal development conditioned by the loss of [r] in onset clusters. It is proposed that the aerodynamic requirements of the apical trill condition a falling fundamental frequency (F0) contour at syllable onset. Specifically, it is hypothesized that the pressure build-up needed to initiate and maintain trilling increases airflow across the glottis, which has the effect of increasing the Bernoulli effect and raising the rate of vocal fold vibration. Evidence from two studies supports this hypothesis.

Much of the literature on tonogenesis has focused on the development of high and low tones after voiceless and voiced onset clusters respectively [6,7,8]. An early explanation of these changes was aerodynamic. Problems with this view [9] led to an alternative hypothesis, namely that F0 differences are due to greater tension on the vocal folds during voiceless than voiced stop productions [10]. We return to an aerodynamic view to explain tonogenesis following the loss of an apical trill. This move is motivated by the stringent aerodynamic requirements of trill production [11] and their potential effect on transglottal airflow. Data from Thai and from non-language trill production are considered.

In the first study, we measured the effect of apical trilling on airflow and F0. Five native Thai speakers were asked to produce nine groups of four monosyllabic words having the same rime and tone but varying in onset from Th to Thr to T to Tr (T = voiceless stop consonant) in a carrier phrase. Results from an analysis of variance indicated that the onset type significantly affected the amount of F0 fall at the beginning of the syllable. Further tests revealed, consistent with the hypothesis, that the Thr onset type had a greater fall than the Th onset type and that the Tr onset type had a greater F0 fall than the T onset type. In the case of the airflow data, we found that [r] had a higher airflow than the vowel onset of the ThV or TV syllable types and that [r] in the Thr context had a greater airflow than in the Tr context. Thus, the greater F0 fall in onsets containing [r] is associated with greater airflow in trill than trill-less onsets.

In the second study, the effect of venting oral air pressure during trill production on F0 was investigated. The hypothesis was that venting air pressure would cause a rise in the volume of air moving across the glottis, and thus raise F0, because additional air in the oral cavity would be needed to maintain the pressure at the apical closure to initiate and maintain trilling. Auditory and airflow recordings of trill produced in [a_a] context with flat intonation by two phonetically trained speakers were made. Vented and non-vented productions were recorded in counterbalanced blocks. The venting was accomplished by introducing a plastic catheter into the post-palatal region of the oral cavity via the buccal sulcus. All trill productions were made with the catheter in place, which allowed venting in one condition, and was pinched to close venting in the other. In the analysis, to normalize for any F0 drift caused by intonational changes, the F0 of the midpoint of the proceeding vowel was subtracted from F0 of each of the trill pulses measured. Results from an analysis of variance indicated, consistent with the hypothesis, that the vented trills had a significantly higher F0 than the non-vented trills.

The results from these studies indicate that a high rate of airflow across the glottis during apical trill production produces a falling F0 contour into the following vowel. This falling

contour can be maintained during and after the loss of a trill, potentially creating a new tonal contour in a language. Thus, the results of the studies reported here reopen the possibility of explaining certain types of tonogenesis in terms of aerodynamic forces.

References

[1] Huffman, F.E. (1967). An outline of Cambodian grammar. PhD Thesis. Cornell University.
[2] Noss, R.B. (1968). The treatment of *[r] in two modern Khmer dialects. In N. Zide (Ed.) Studies in Comparative Austroasiatic Linguistics, The Hague: Mouton, pp. 89-95.
[3] Pisitpanporn, N. (1999). A note on colloquial Phnom Penh Khmer. In G. Thurgood, (Ed.) Papers from the Ninth Annual Meeting of the Southeast Asian Linguistics Society, Tempe, Arizona: Program for Southeast Asian Studies, Arizona State University, pp. 243-248. [4] Thach, 1999;
[5] Wayland and Guion (under review). Sound changes following the loss of /r/ in Khmer: A new tonogenetic mechanism?
[6] Haudricourt, A.-G. (1954). De l'origine des tons en Viêtnamien. Journal Asiatique 242, 69-82. [7] Matisoff 1973,
[8] Hombert, J.-M. (1978) Consonant types, vowel quality, and tone. In V. Fromkin (Ed.) Tone: A Linguistic Survey, NY: Academic Press, pp. 77-111.
[9] Ohala, J.J. (1978). Production of tone. In V. Fromkin (Ed.) Tone: A Linguistic Survey, New York: Academic Press, pp.5-39.
[10] Löfqvist, A., Baer, T., McGarr, N.S., & Story, R.S. (1989). The cricothyroid muscle in voicing control. Journal of the Acoustical Society of America 85, 1314-1321.
[11] Solé, M-.J. (2002) Aerodynamic characteristics of trills and phonological patterning. Journal of Phonetics 30, 655-688.

Mechanisms of sound change in Romance: From gemination to degemination in Italy

John Hajek & Mary Stevens
University of Melbourne

The historical contrast between short and long consonants found in Latin survives today in almost all of Central and Southern Italy, eg Tuscan /sete/ 'seven' v /sette/ 'thirst' but has been lost as a result of a subsequent process of degemination in Northern Italy and the rest of the Romance-speaking world, eg Lombard /sed/ v. /set/. Whilst there has been much discussion in the phonological literature that tries to describe and account for the shift, little interest has been shown in trying to understand the phonetic basis, if any, for the phenomenon. The aim of this paper is to try to provide such a phonetic account, drawing on the results of our acoustic investigation of geminate voiceless stops /pp tt kk/ in Sienese Italian, a variety of Tuscan spoken in Central Italy. Our experimental results, we argue, allow us to formulate an entirely new model of the mechanism of sound change involved in degemination.

A detailed analysis of our spontaneous speech data (4 Ss) has uncovered a previously unreported process of phonetic degemination, associated with a series of glottalization effects in vowel offsets before /C:/. These glottalization effects, previously not known or only barely so, present much variation, ranging from creak and glottal stop through to preaspiration, frication and breathy voice. When they occur, they trigger significant phonetic shortening of the otherwise expected long consonant. While we did not find glottalization effects to be conditioned by stress, we did find them to be predictably affected by differences in vowel height. We argue that this interaction is a significant factor in motivating the phonological process of degemination as a gradual, conditioned sound change over time, as found in some dialects today, such as in Borgo San Sepolcro (C. Italy) and Bolognese (N. Italy) where optional breathy voice has already been associated with degemination and secondary overlength, e.g. Latin /panna/ > [pa::na] ~ pa:hna] 'cream' (Hajek 2000).

While in Sienese Italian, there are good arguments to support the view that preaspiration and the other glottal effects should still be included in the overall consonant duration (and not part of the vowel) we also show how perceptual reinterpretation a' la Ohala's listener oriented model of sound change can lead to the progressive reassignment of the glottalized portion to the vowel, and in degeminating dialects to the progressive loss of consonant gemination over time. Such an account relies on consistent perceptual evidence that shows Italian speakers rely heavily on vowel duration to identify consonant length (eg Rochet et al 1995 and others).

References

Hajek, J (2000) How many moras? Overlength and maximal moraicity in Italy. In Repetti, L (ed.) Phonological theory and the dialects of Italy. John Benjamins, Amsterdam/Philadelphia:111-135

Rochet, B. L., and Rochet, A.P. (1995). The perception of the singlegeminate consonant contrast by native speakers of Italian andanglophones. Proceedings: ICPhS 95, Stockholm, Vol. 3, pp. 616-9.

A vowel height split explained: Compensatory listening and speaker control

Carlos Gussenhoven
University of Nijmegen

Two experiments were conducted which show that high vowels [i,y,u] sound longer than low vowels [E, OE, O] when acoustic durations are equal. The effect is comparable to a 25 ms difference in acoustic duration. The stimuli were natural vowels whose acoustic durations had been manipulated with the help of Praat and which were judged for duration by listeners with different language backgrounds. The following two points will be argued for:

1. The finding results from 'compensatory listening' (Pierrehumbert 1979, Silverman 1987). Because low vowels are intrinsically longer, owing to the larger distance the tongue body travels from a consonantal contact to the vowel target, listeners subtract some of the 'extra' time from low vowels, causing high vowels to sound longer. Earlier, Pierrehumbert showed that the intrinsic advantage of earlier pitch peaks was similarly compensated for, since listeners hear them as lower-pitched than later peaks when F0's are equal. Silverman showed that high vowels like [i:], which have higher intrinsic F0, are heard as lower-pitched (less prominent) than low vowels like [a:] when F0's are equal.

2. The finding explains a vowel split that took place in Dutch dialects spoken in Belgium in which duration is used as an enhancement of a tone contrast. The more conservative dialects have a tonal distinction between 'Accent 1' and 'Accent 2', which occurs on segmentally identical rhymes of main stressed syllables. Accent 2 not only differs from Accent 1 in F0, usually by having a later peak, but also in having a longer duration, particularly in IP-final position. Speakers of innovating dialects have a third way of improving the discriminability of the contrast: to boost the duration effect, they have closer vowels for Accent 2 than for Accent 1. Data from the dialect of Mechelen-aan-de-Maas are shown in (1) (Verstegen 1996). Here, the tone contrast still exists. In other cases, the tone contrast was given up after similar splits. In general, dialects that have lost the tonal contrast tend to have larger numbers of tone-induced vowel splits than dialects in which the tone contrast still exists.

(1)	Accent 1	Accent 2
	GEEl 'yellow-ATTR'	Geel 'yellow-PRED'
	wEEx 'road-PL'	weex 'road-SG'
	GOOn 'go-1SG,PRES'	Goon 'go-1PL,PRES'
	nOOl 'needle-SG'	nool@ 'needle-PL'

This shows that 'compensatory listening' needs to be presented in the wider context of speaker control in speech production (Kingston & Diehl 1994). Belgian Dutch vowel height differentiation is one of the many ways in which speakers have recruited phonetic resources to support what they felt was a precarious phonological contrast. In addition to height differentiation as a way of boosting the durational difference, closing diphthongs were split into diphthongs when co-occurring with Accent 2 and vowel+glide combinations when co-ocurring with Accent 1. This again is to be explained as a way of creating a shorter vowel duration in syllables with Accent 1, since unlike the second element of the diphthong, the glide is not included in the vowel percept. A third experiment will be reported that shows this to be case for the same group of Dutch listeners as took part in the two experiments that tested for vowel height

effects. The Dutch tonal dialects present an interestingly different set of interactions between tone and other phonetic parameters from Asian languages.

References

Kingston, J. & Diehl, R., 1994, Phonetic knowledge. Language 70, 419-454.

Pierrehumbert, J.B., 1979. The perception of fundamental frequency declination. JASA, 66, 363-369.

Silverman, K.E.A., 1987. The structure and processing of fundamental frequency contours. PhD Thesis, Cambridge University.

Verstegen, V., 1996. Bijdrage tot de tonologie van Oostlimburgse dialecten. Een eeuw Limburgse dialectologie. H. van de Wijngaard (ed.), Hasselt/Maastricht: VLDN/Vereniging Veldeke Limburg, 229-234.

Intonational overload: Uses of the H* !H* L- L% contour in read and spontaneous speech

Julia Hirschberg*, Agus Gravano*, Ani Nenkova*, Elisa Sneed† & Gregory Ward†
Columbia University and Northwestern University†*

Intonational contours are OVERLOADED, conveying different meanings in different contexts. In this paper we examine two potential uses of the DOWNSTEPPED contours in Standard American English, in the Boston Directions Corpus of read and spontaneous speech. We investigate speakers' use of these contours in conveying discourse topic structure and in signaling GIVEN vs. NEW information and the relationship between these two functions.

Exemplar-based phonology and the time problem: a new representational technique

Robert Kirchner
University of Alberta

A growing body of phonetic and phonological research motivates a speech processing system in which words are represented as cohorts of exemplars – detailed memories of the sound signal (and articulatory plan) for previous tokens with the same semantic label [1]-[7]. To summarize the thrust of this research, such a model naturally accounts for a range of stochastic and frequency-sensitive phonetic and phonological patterns, sound changes, acquisition phenomena, and experimental results – all of which are beyond the capacity of standard, symbolic approaches to phonology. Furthermore, the widely attested development of categorical phonological patterns from gradient phonetic variation follows from the system's inherent tendency to sharpen existing patterns.

Exemplar-based classification, however, presupposes that each exemplar can be assigned to a particular point in some fixed-dimensional mathematical space [7], such that it has a measurable distance from all other exemplars. But the speech signal is inherently of variable length. Indeed, the time dimension problem has likewise prevented a range of pattern recognition techniques from being successfully applied to speech processing, including neural nets and support vector machines, as these also require a fixed-dimensional input.

A new technique, the string kernel [8], [9] extended to speech recognition by Goddard et al. [10], permits variable-length signals to be mapped to fixed-dimensional space, by examining whether they contain the same sub-sequences. Specifically, one takes the Euclidean distance of the inner products (a sort of count) of sub-sequences (usually 2 to 4 elements) in the original strings. Goddard et al, report that on a digit recognition task, their kernel-based classifier outperforms a hidden Markov model, the standard of the speech recognition industry.

In this paper, I present several modelling experiments, testing the feasibility of this technique as applied to an exemplar-based speech processing model, implemented in Matlab. The first experiment shows that on a similar digit recognition task, a string kernel which counts non-contiguous as well as contiguous sub-sequences (with a weight that exponentially decays with the distance between the elements) performs even better than the 'binary" kernel used by Goddard et al. The second experiment (currently in progress) tests whether the relative distance scores computed by the kernel-based k-nearest-neighbours classifier correspond to a priori notions of distance on a range of phonetic dimensions (e.g. tokens of [bi:k] are closer to [bi:g] than to [bi:b] or [beɪg]). The third experiment addresses the reverse mapping problem: in order serve as the basis for a full speech-processing model, including production, the string kernel must be reversible, i.e. capable of mapping back to an approximation of the original signal. I propose an algorithm which reconstructs the signal from the kernel representation, by recursively factoring the inner products into floors and remainders. This algorithm, applied to the spoken digit corpus of the first experiment, matches the original string with 99% accuracy. Using this algorithm, the model can synthesize recognizable outputs, based on centroid properties of the target word cohort, plus the influence of strongly instantiated patterns from similar exemplars.

It thus appears that the string kernel technique provides a feasible, general solution to the time problem in exemplar-based phonology. With this global comparison method, exemplar-based models need not appeal to phone units (nor other sub-word phonological units) for local comparison, as in, e.g., [7]. Rather, the (quasi-)phonemic organization of speech can be shown to

emerge from the relatively stable correlations between articulatory gestures and perceptual cues over the lexicon of exemplars, cf. [11]. The problem of segmentation of connected speech into words remains an issue for future research.

References

[1] Johnson, K. (1997) Speech perception without speaker normalization. In K. Johnson & J. Mullennix (eds.) Talker variability in speech processing. San Diego: Academic Press.

[2] Jusczyk, P. (1997) The discovery of spoken language, Cambridge, MA: MIT Press.

[3] Kirchner, R. (1999) Preliminary Thoughts on 'Phonologization' within an Exemplar-Based Speech Processing System. UCLA Working Papers in Linguistics [ROA 320-699].

[4] Bybee, J. (2001) Phonology and language use, Cambridge: Cambridge University Press.

[5] Kirchner, R. & C. Fraser (2002) Modelling phonologization in an exemplar-based lexicon, poster presented at 3rd International Mental Lexicon Conference, Banff, Alberta, Oct. 6, 2002.

[6] Pierrehumbert, J (2001) Exemplar dynamics: word frequency, lenition, and contrast, in J. Bybee, and P. Hopper (eds.), Frequency effects and the emergence of linguistic structure, 137-157 Amsterdam: John Benjamins.

[7] Pierrehumbert, J. (2002) Word-specific phonetics, in C. Gussenhoven & N. Warner (eds.), Laboratory Phonology VII, Berlin: Mouton de Gruyter. 101-140.

[8] Watkins, C. (1999) Kernels from matching operations, Computer Science Department Technical Report 98-07, Royal Holloway, U London, 1999.

[9] Lodhi, H., C. Saunders, J. Shawe-Taylor & N. Cristianini (2001) Text classification using string kernels, Journal of Machine Learning Research 2:419-444.

[10] Goddard, J., A. Martinez, F. Martinez, H. Rufiner, A comparison of string kernels and discrete hidden Markov models on a Spanish digit recognition task, JASA 112:5, pt. 2, 2002.

[11] Beckman, M., and J. Pierrehumbert (in press) Interpreting 'phonetic interpretation' over the lexicon, in J. Local, R. Ogden, & R. Temple (eds.) Papers in Laboratory Phonology VI, Cambridge: Cambridge University Press.

A Competitve, Coupled Oscillator Model of Moraic Structure

Hosung Nam
Yale University and Haskins Laboratory

This paper discusses how to interpret and implement the mora as a timing unit in the articulatory phonology framework and focuses on 1) the shortening effect of VC sharing a mora and 2) the asymmetry of duration between a shared mora and an unshared mora. We argue that a single mora is equivalent to the sequential coordination of a pair of gestures in coupling graphs (gestural coordination structures in dynamical terms) when this sequential coordination is not modulated by a competing coordination relation. We extend Nam and Saltzman's coupling graphs [2], for which intergestural phonetic variations in timing shown at syllable edges and across boundaries are accounted by allowing multiple and competitive coupling of articulatory gestures modeled as oscillators. In that model, the weightlessness of onset consonants (lack of mora) emerges from the synchronous nature of C-V coordination and from the competition between C-C and C-V coordination in onsets.

To extend this model to the problem of coda mora-sharing, we propose splitting a consonant gesture into two gestures (closure (CLO) and release (REL)) [3], which need to be specifically coordinated with respect to one another We hypothesize that in an unshared mora case (e.g. vowel and consonant having their own mora), the preceding vowel is dynamically coordinated with the following CLO and the CLO is coupled to the following REL. Hence, each mora corresponds to each dynamic coordination: V-CLO and CLO-REL. On the contrary, in a shared mora case (e.g. vowel and consonant sharing a mora), an additional coordination (V-REL) is added in a manner in which V-REL (one mora) is in competition with V-CLO and CLO-REL (two moras). The simulation results showed that when the oscillators are parameterized identically, the resultant VC duration in the shared mora case generated by this competitive coupling structure model is shorter than the VC duration of two unshared moras but longer than a single segment (C or V) duration of an unshared mora.

A recent study has demonstrated that distinctive moraic representation has a correlation with segment duration as phonetic realization [1]. It has been argued that a coda consonant has its own mora in Hindi but shares it with the preceding vowel in Malayalam. In Levantine Arabic, the mora sharing depends on the length of a vowel: a coda consonant shares a mora with the preceding long vowel while a coda consonant following a short vowel has its own mora. As consistent with their prediction, due to V2 and C sharing a mora in V1V2C, V1V2 exhibits shorter duration than both VC and VV where each segment has its own mora, and also C in V1V2C is shorter than C in VC.

However, their shared-mora representation fails to account for fact that the duration of the entire V1V2C shared mora sequence is longer than other kinds of two mora sequences in Levantine Arabic. The durations (199.2, 193.1, 164.7) of V1V2C, in which V1 has its own mora and V2 and C share a mora, are longer than those of VV (161, 123.9, 114.2) and VC (168.3, 179.4, 148.6), both of which also have two moras (the numbers in the parentheses indicate the mean durations (ms) in Jordanian, Syrian and Lebanese Arabic, respectively). Such durational effects do emerge from our coupling model as a result of competition between coordination specifications.

References

[1] Broselow, E., Chen, S., and Huffman, M. 1997. "Syllable weight: convergence of phonology and. Phonetics", Phonology 14, pp. 47-82

[2] Nam, H. & Saltzman, E. 2003. "A competitive, coupled oscillator of syllable structure," Proceedings of the XIIth International Congress of Phonetic Sciences, Barcelona, 3-9 Aug 2003. Vol. 3, pp. 2253-2256.

[3] Steriade, D. 1993. "Closure, release, and nasal contours," Nasals, nasalization, and the velum, The phonetics of nasal phonology: theorems and data, M. K. Huffman & R. A. Krakow (eds.), [Phonetics and Phonology Series, Vol. 5] San Diego, CA: Academic Press. pp. 401-470.

Tracking the timecourse of multiple context effects in assimilated speech

Bob McMurray & David Gow
University of Rochester

Recent research has demonstrated the sensitivity of lexical activation to fine-grained phonetic detail (Andruski, Blumstein & Burton, 1994; Gow & Gordon, 1995; McMurray, Tanenhaus & Aslin, 2002). However, few studies have systematically examined how such detail is integrated over time in word recognition. Place assimilation is an ideal context in which to examine these issues. In English, segments with underlying coronal place often show partial assimilation to the place of a subsequent labial or velar (e.g. /m/ or /k/). This produces a surface form with acoustic and articulatory properties that are partially consistent with both coronal and non-coronal place (e.g. Green Boat => Green$_m$ Boat). Listeners can use the resulting acoustic/phonetic cues to *anticipate* upcoming material and use post assimilation context to *resolve* the ambiguous consonant (Gow, 2001; Gaskell & Marslen-Wilson; 1998). However, there has been no detailed examination of the timecourse of these effects to probe their relationship to one another or to lexical activation. Thus, we used an eye-tracking task in which subjects respond to spoken instructions to manipulate visual objects whose names have linguistic properties of interest (Tanenhaus, Spivey-Knowlton, Sedivy & Eberhart, 1995). Fixation probabilities in this task mirror the activation of lexical items as speech unfolds over time (Allopenna, Magnuson & Tanenhaus, 1998).

Experiment 1 examined anticipation. 21 subjects heard phrases such as "Select the maroon goose" in which "maroon" was either naturally assimilated or unassimilated. Subjects selected (with a mouse) the picture from a screen containing all four combinations of word-final coronal/non and word-initial coronal/non: a maroon goose, a maroon duck, a patriotic goose and a patriotic duck. More eye-movements to the target (maroon goose) were observed immediately after an assimilated than an unassimilated coronal ($t(20)=2.13$, $p=.046$) and their were fewer eye-movements to the competitor, maroon duck ($t(20)=2.43$, $p=.024$). These results suggest that assimilation can help prepare the system by increasing priors for potential non-coronal targets, and decreasing them for potential coronal competitors.

Assimilation can often create lexical ambiguity: "cat" is ambiguous between "cat" and "cap", when it assimilates bilabial place (Cat Box => Cat$_p$ Box). This ambiguity can be resolved when the post-assimilation context is integrated with the subphonemic cues for assimilation in the ambiguous consonant. In Experiment 2 subjects heard phrases like "Select the cat box", in which cat was either assimilated or not. Screens contained pictures of "cat box" (coronal-non), "cat drawing" (coronal-coronal), "cap box" (non-non) and "cap drawing" (non-coronal). Eye-movements were analyzed for looks to the item corresponding to the *initial* coronal (cat vs. cap). *Late* in the trial, more looks were seen to the initial coronal (cat) than the non-coronal (cap) when the second item was non-coronal (box). The reverse was found ($F(1,20)=11.2$, $p=.003$) when the second item was a coronal (drawing). This supports a regressive ambiguity resolution mechanism consistent with feature-parsing (Gow, 2003). Additionally, *anticipatory* effects on *early* eye-movements were marginally significant for the correct target, e.g. "box" ($t(20)=1.84$, $p=.081$), and significant for looks to the competitor ($t(20)=2.45$, $p=.024$), suggesting both anticipatory and regressive mechanisms may be active for the same items (though not necessarily simultaneously).

These experiments support the on-line use of subphonemic detail for anticipating upcoming material and resolving ambiguity and suggest that the eye movement paradigm may be ideal for studying the integration of acoustic detail over time. We argue for active recognition process that is sensitive to phonological regularities in the signal and in which information integrated over a relatively large temporal range.

References

Allopenna, P., Magnuson, J.S. and Tanenhaus, M.K. (1998) Tracking the time course of spoken word recognition using eye-movements: evidence for continuous mapping models. Journal of Memory and Language, 38(4), 419-439.

Andruski, J.E., Blumstein, S.E. and Burton, M.W. (1994) The effect of subphonetic differences on lexical access. Cognition, 52, 163-187.

Gaskell, G. and Marslen-Wilson, W. (1998) Mechanisms of phonological inference in speech perception. Journal of Experimental Psychology: Human Perception and Performance, 24, 380-396.

Gow, D. (2001) Assimilation and Anticipation in continuous spoken word recognition. Journal of Memory and Language, 45, 133-139.

Gow, D. (2003) Feature parsing: Feature cue mapping in spoken word recognition. Perception & Psychophysics, 65(4), 575-590.

Gow, D. and Gordon, P. (1995) Lexical and prelexical influences on word segmentation: Evidence from priming. Journal of Experimental Psychology: Human Perception and Performance, 21(2), 344-359.

McMurray, B., Tanenhaus, M.K. and Aslin, R.N. (2002) Gradient effects of within-category phonetic variation on lexical access. Cognition, 86(2), B33-B42.

Tanenhaus, M.K., Spivey, M.J., Eberhard, K.M. and Sedivy, J.C. (1995) Integration of visual and linguistic information in spoken language comprehension. Science, 268, 1632-1634.

Lexical Access and Hyperarticulation

Benjamin Munson
University of Minnesota

A growing body of literature has documented the existence of word-specific phonetic patterns. Instances of the 'same' phonological category may differ systematically across the words in which they are attested, as a function of a variety of factors. For example, English speakers hyperarticulate vowels in words with high phonological neighborhood densities (ND), relative to the same vowels in words with low ND (Wright, in press). That is, the acoustic vowel spaces associated with high-ND words are more dispersed than those for low-ND words. One explanation for this density-dependent dispersion effect (henceforth, the 3D effect) is suggested by the work of Jurafsky et al. (2003). Jurafsky et al. hypothesized that difficulties in lexical access are related to hyperarticulation: when normal speakers encounter difficulty in lexical access, they tend to hyperarticulate the item they have had difficulty accessing. This project examines the hypothesis that difficulties in lexical access associated with high-ND words underlie the 3D effect. To address this hypothesis, we examined vowel articulation in words that varied both in frequency and in ND, in two experimental conditions. In the first condition (the immediate-response condition), participants were presented with a word on a computer monitor and were instructed to read it as quickly as possible. In the second condition (the delayed-response condition), participants were presented with the same words, but a 1000 ms delay interval was enforced between presentation of the word and participants' response. As in previous work (Balota & Chumbley, 1985), we presume that production in the immediate-response condition is initiated before lexical access has been fully completed. Responses in that condition reflect the influence of lexical access on speech production. Productions in the delayed-response condition do not reflect the effects of lexical access, as they are presumed to occur after this has been completed. Fifteen speakers participated in the experiment. All were native adult speakers of English with no reported history of speech, language, or hearing disorders. The stimuli were 80 high and low frequency, high and low ND CVC words. There were 20 words for each frequency/ND combination. The four lists were counter-balanced for vowel quality and final consonant voicing and manner. Speakers produced the 80 words in both the immediate- and delayed-response conditions. The first and second formant frequencies of each vowel were measured. Vowel-space dispersion was calculated using a method from Bradlow et al. (1996). Dispersion was calculated separately for the four lists of words in both the immediate-response and long-delay conditions. A three-factor within-subjects ANOVA found significant main effects of frequency and ND. As in Wright (in press), vowel spaces associated with high-ND words were more dispersed than those associated with low-ND words. The effect of frequency arose because the vowel spaces associated with low-frequency words were more dispersed than those associated with high-frequency words. In addition, the three-way interaction between delay condition, frequency, and ND was significant. Post-hoc tests of significant main effects found that the effect of ND was present in both the immediate- and delayed-response conditions. The effect of frequency on dispersion, however, was not present in the delayed-response condition. Vowel durations did not differ among the conditions, suggesting that this was not a confounding factor. Results do not support the hypothesis that the 3D effect is a consequence of difficulties in accessing words residing in dense phonological neighborhoods. Rather, it suggests that the 3D effect represents an active attempt to maximize the clarity of

perceptually difficult words, as proposed by Wright (in press). Alternatively, the 3D effect may reflect speakers' selective encoding of clearly produced exemplars of the easily confusable high-ND words, as proposed by Pierrehumbert (2003). The effect of frequency on vowel-space dispersion, however, may be related to difficulties in lexical access, as this effect was not present in the long-delay condition.

References

Balota, D., & Chumbley, J. (1985). The locus of word-frequency effects in the pronunciation task. *Journal of Memory and Language, 24*, 89-106.

Bradlow, A., Torretta, G., & Pisoni, D. (1996). Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication, 20*, 255-272.

Jurafsky, D., Bell, A., & Girand, C. (2003). The role of the lemma in form variation. In N. Warner and C. Gussenhoven (Eds). *Papers in Laboratory Phonology VII* (p. 1-34). New York: Mouton de Gruyter.

Pierrehumbert, J. (2003). Word-specific phonetics. In N. Warner and C. Gussenhoven (Eds). *Papers in Laboratory Phonology VII* (p. 101-139). New York: Mouton de Gruyter.

Wright, R. (in press). Factors of lexical competition in vowel articulation. In J. Local, J., R. Ogden, and R. Temple (eds.). *Papers in Laboratory Phonology VI.* Cambridge: Cambridge University Press.

Language Contact in Bilingualism: Phonetic System Interactions

James E Flege
University of Alabama

The sound systems of two languages influence one another when they enter into contact, primarily when people become bilingual. This contribution will describe research carried out within the framework of the Speech Learning Model (SLM). The SLM proposes two mechanisms for inter-system phonetic change in bilingualism. The vowels and consonants (sounds) used to distinguish words in the native (L1) and second language (L2) may come to resemble one another through phonetic category assimilation. This happens when new categories are not established for L2 sounds (either because L1 categories are too robust when L2 learning begins, or insufficient input has been received). When new phonetic categories are established for L2 sounds, L2 sounds might dissimilate from native-speakers norms. This permits bilinguals to maintain contrast between new L2 sounds and pre-existent L1 sounds. In either case, bilingual speakers of an L2 might differ from monolingual speakers of the target L2. The two proposed mechanisms will be illustrated from the results of studies examining the production and perception of L2 vowels and consonants. Work showing the influence of L2 learning on performance in the L1 will be also presented, along with evidence of intra-system reorganization among elements making up the L1 phonetic system.

Perceptual learning and plasticity in L2 learners: building a new system for phonological processes

Isabelle Darcy*†, Sharon Peperkamp*‡ & Emmanuel Dupoux*
Laboratoire de Sciences Cognitives et Psycholinguistique, EHESS - ENS – CNRS, Paris, Johannes Gutenberg Universität Mainz† and Université de Paris 8‡*

The word recognition system copes easily with the great variability of spoken language. This study explored the perception in French and English of variability in word-final consonants conditioned by the first consonant of the next word. In this context, French has a tendency towards voicing assimilation but not place assimilation. English has a tendency towards place but not voicing assimilation. Subjects were 56 late French English bilinguals, whose L1 was French or English. A word detection task provided evidence that fluent bilinguals had acquired the ability to compensate perceptually for the assimilatory pattern of their L2, without compromising their processing of L1.

Comparable stimulus sets were developed for both French and English. Each set of sentences provided examples of the English place assimilation process (on coronal stops) and of the French voicing assimilation process (on obstruents). Both French and English sentences entail voicing assimilation occurrences, a native or non-native process respectively. Similarly, French and English sentences contain place assimilation occurrences, a non native or native process respectively. In the sentence, targets occurred either without any change (baseline) or with a one-feature change on their last consonant, in three different contexts. Changes correspond to voicing or place assimilation, made viable or non viable by the context (i.e. next word's initial), which triggers or not assimilation: Non trigger contexts are for both contrasts sonorants or coronal fricatives, paired to the baseline ('bla[k+r]ag' 'fa[t+m]onkey') and the non-viable change condition ('bla[g+l]ine', 'fa[p+s]quirrel'). Trigger contexts for voicing are obstruent consonants agreeing in voicing ('bla[g+g]love'), and for place, they agree in place (labial or velar, 'fa[p+p]uppy'). Measured are word detection rates in each of the three conditions (viable, non viable and identical), as a function of type of contrast (place or voicing). The full stimulus set of 288 test items contained 32 target words in 3 contexts each (one per condition). Subjects heard each target in all three contexts, but in different sentence frames, couterbalanced across three experimental lists. Sentences were recorded by trained female native speakers; changes were deliberate, in order to be neutralizing. Clarity of stimuli was validated by a separate experiment in a forced choice task on the last consonant of the excised target words, with other 18 native speakers in each language. In isolation, changes were perceptually treated as tokens of the intended phoneme, i.e. the voicing or place counterpart of the underlying phoneme, equally in both change conditions. 56 subjects tested in L1 (Darcy et al. under review) interpret such a contextual change with reference to the context: they compensate for assimilation in legal contexts (~65% detection for viable change over ~20% in non viable, $p<.001$) – for the native assimilation type, not for the non-native one (equal in both conditions), demonstrating prior conditioning by language specific phonological knowledge.

The present study assessed the extent to which lexical access in L2 reflects implicit learning of the L2 assimilation pattern, versus intrusion from the (incorrect) L1 assimilation pattern (like for non-native contrasts, e.g. Best 1995). The subjects were the same 56 late learners of L2 (30 French and 26 Americans), with no or very little exposure prior to school (11-12 yrs.). Participants whose amount of native exposure to L2 (i.e. living in the country) is

comprised between 0 and 24 months are 'beginners'. 'Advanced' participants all have had at least 2 years of native exposure. Beginners show the same compensation pattern for L1 and L2: they interpret a change in L2 (e.g. voicing) according to L1-positional categorization patterns, not according to the native or non native status of the change in the language they hear. In contrast, advanced learners developed the capacity to categorize changes according to the context differently in L1 and L2. In conclusion, the word recognition system appropriately compensates for phonological variation involving an abstract language-specific phonological knowledge which can be adapted to a new language. Learning of these positional categorization patterns is possible perceptually within a few years of exposure to L2; this new system is independent of the first one, as no interference between L1 and L2 has been observed.

References

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

Darcy, I., Ramus, F., Christophe, A., Kinzler, K. & Dupoux, E. (under review) Language specific effects in compensation for phonological variation. Journal of Memory and Language.

Consonant Timing in L2 English: The Emergence of a Default Pattern

Mariko Yanagawa
Yale University

It is known that in English, two consonant gestures across a boundary may overlap substantially in time (e.g. Browman and Goldstein 1990, Byrd 1996, Zsiga 1994). In Russian, in contrast, the two gestures are separated by a lag with no overlap (Zsiga 2000, Kochetov 2002). Zsiga (2003) examined whether these different patterns were acquired by L2 learners of English and Russian, and found that English learners of L2 Russian were able to imitate the Russian-type "no-overlap" pattern, while Russian learners of L2 English had greater difficulty imitating the English-type "overlap" pattern. L2 acquisition often reveals patterns that are independent of both L1 and L2, and are crucially unmarked (Broselow et al. 1998). One possible interpretation of Zsiga's results is that the "no-overlap" pattern is more unmarked, and possibly easier to learn, as native speakers of both languages appeared to employ this pattern in L2.

The present study aims to investigate the possible effect of L1 cross-word overlap pattern on L2 overlap by employing languages with a wide variety of L1 overlap patterns. There are two possibilities: (1) L2 overlap patterns vary as a function of L1 overlap; or (2) L1 overlap has no effect on L2 overlap, and a default overlap pattern with little or no overlap always emerges in L2. An experiment was conducted to test overlap in L2 English consonant sequences as well as L1 consonant sequences produced by speakers of different L1 backgrounds, i.e. Cantonese, Taiwanese, German, French, and Japanese. Overlap was measured acoustically using Zsiga's (2000, 2003) method.

Acoustic analysis revealed that the speakers all exhibited the "no-overlap" pattern in L2 English, which is not manifested in their L1 nor in native English. For example, while Cantonese speakers showed even more extensive overlap than English in L1, the same speakers did not carry this native pattern into the comparable clusters in L2 English. The results suggested that an unmarked, default overlap pattern emerges in L2.

The emergence of a default overlap pattern in L2 can be explained by preferred modes of coordination (Browman and Goldstein, in preparation). Two intrinsic modes of coordination, i.e. synchronous (in-phase) and sequential (out-of-phase), have been proposed to be intrinsically available without specific learning (Saltzman 1995). We hypothesize that learners opt for the sequential mode rather than the synchronous mode, because the former, which is the "no-overlap" pattern in speech, allows learners to have more time in planning and/or producing sequential L2 words.

References

Broselow, Ellen, Chen, Sui-I, and Wang, Chilin. (1998) "The Emergence of the Unmarked in Second Language Phonology", Studies in Second Language Learning 20:261-280.

Browman, Catherine P. and Goldstein, Louis. (1990) "Tiers in Articulatory Phonology, with Some Implications for Casual Speech", in J. Kingston and M. E. Beckman (eds.) Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech. Cambridge, UK: Cambridge University Press.

Browman, Catherine P. and Goldstein, Louis. (in preparation) Articulatory Phonology. Cambridge, UK: Cambridge University Press.

Byrd, Dani. (1996) "Influences on Articulatory Timing in Consonant Sequences", Journal of Phonetics 24:209-244.

Kochetov, Alexei. (2002) Production, Perception, and Emergent Phonotactic Patterns: A Case of Contrastive Palatalization. London and New York: Routledge

Salzman, Elliot. (1995) "Intergestural Timing in Speech Production: Data and Modelling", in K. Elenius and P. Branderud (eds.) Proceedings of the XIIIth International Congress of Phonetic Sciences. Stockholm:KTH and Stockholm University.

Zsiga, Elizabeth (1994) "Acoustic Evidence for Gestural Overlap in Consonant Sequences", Journal of Phonetics 22:121-140.

Zsiga, Elizabeth (2000) "Phonetic Alignment Constraints: Consonant Overlap and Palatalization in English and Russian", Journal of Phonetics 28:69-102.

Zsiga, Elizabeth. (2003) "Articulatory Timing in a Second Language", Studies in Second Language Acquisition 25: 399-432.

Variation in vowel production by English-Arabic bilinguals

Ghada Khattab
University of Newcastle

This study presents a phonetic analysis of vowel production by three balanced English-Arabic bilinguals, aged between 5 and 10 and living in Yorkshire, England. The study departs from traditional investigations of bilingual phonological acquisition in several ways. First, a broader definition than usual is assumed of what is meant by 'phonological acquisition'. Most studies adopt a broadly phonemic approach, with a focus on the acquisition of features relevant for lexical contrast in the two languages. Such an approach, however, marginalises those systematically variable aspects of production which may not be relevant for lexical contrast but which carry social-indexical information (linked to speaker sex, age, class etc.). The learning of sociophonetic features constitutes an essential part of the acquisition process for monolinguals (Docherty et al in press). In the case of bilinguals, sociophonetic variation is even more pervasive: the child's linguistic input may comprise standard, non-standard, and non-native varieties for two languages. Few studies have considered the sociophonetic dimensions of acquisition, or their impact on the cognitive representation of two languages.

In light of the interest in sociophonetic variation, a different methodology was adopted to establish the relevant parameters of variation for the children. Bilingual studies tend to rely on generalised descriptions of the target phonologies (often based on standard varieties of the two languages). Differences between bilinguals' production and the assumed targets are often therefore interpreted as errors, imperfect learning, or interference between phonological systems. In the current study, a detailed picture was obtained of the sociophonetic features in the children's input. This was achieved by recording age-matched monolingual friends of the bilinguals and the parents of all the children. A total of 23 subjects was therefore recorded.

The methodology furthermore controlled for language mode (Grosjean, 1998), which few phonological studies have done. The choice of language, topic, and interlocutor may all affect the bilingual's choice of linguistic variants. Knowledge of these choices must therefore constitute part of the sociolinguistic repertoire which children must acquire. Separate recording sessions were carried out to assess the bilinguals' interaction with monolinguals and bilinguals, and code-switched portions were analysed separately.

This paper concentrates on six vocalic variables, namely the vowels of the BATH, STRUT, FACE, PALM, START, and GOAT lexical sets (Wells, 1982). These were chosen because of their context- and dialect-specific realisations in Yorkshire English. The results show that sociophonetic aspects are being acquired simultaneously with reflexes of the contrastive system. The bilinguals display a similar range of realisations as the monolingual controls, showing that they have acquired patterns appropriate to their community. However, the findings also reflect the bilinguals' wider linguistic repertoire. This showed in a number of realisations that were exclusive to the bilinguals. Some of these appear to be the result of influence from the bilingual parents' non-native accent, e.g. a close monophthong [o:] in GOAT. However, these were only noted when the children were in a bilingual mode, code-switching into English from an Arabic base. A typical realisation of *boat*, for example, might be [bout] in the children's monolingual mode English, but [bo:t] when inserted in code-switches. Similar effects were seen with consonantal variables, e.g. /r/ being tapped.

This finding constitutes a challenge to Chambers' (2003) suggestion that bilinguals have an 'accent filter' which enables them to eliminate foreign-accent features in the input that they receive. It further suggests that variation in bilinguals' speech production need not result from uncontrolled interference between phonological systems: instead, bilinguals as young as 5 can actively exploit variable phonological resources for specific communicative purposes. This implies that the cognitive storage of two phonological systems may indeed overlap, as is often claimed in bilingual research. But crucially, the degree and effects of overlap may be under speakers' control as part of their phonological knowledge.

References

Chambers, J.K. (2003) Sociolinguistic Theory. Oxford: Blackwell.

Docherty, G., Foulkes, P., Tillotson, J. and Watt, D. (in press) On the scope of phonological learning: issues arising from socially-structured variation. To appear in Laboratory Phonology 8, ed. C. Best and L. Goldstein. The Hague: Mouton de Gruyter.

Grosjean, F. (1998) Studying bilinguals: methodological and conceptual issues. Bilingualism: Language and Cognition 1: 131-149.

Wells, J.C. (1982) Accents of English (3 vols.). Cambridge: Cambridge University Press.

LabPhon 9
Change in Phonology

Ninth Conference on Laboratory Phonology
Thursday 24 June - Saturday 26 June, 2004

US Minority Students Participation Grant

The organizers of the ***9th Conference on Laboratory Phonology (LabPhon 9)*** are pleased to announce the availability of two small grants to support the participation of US minority students in LabPhon 9. These grants are intended to cover registration and travel expenses. To qualify, applicants must be citizens or permanent residents of the USA currently enrolled as students at a US institution and members of an underrepresented minority. Priority will be given to qualified candidates who

- can demonstrate insufficient financial support from their home institution to attend LabPhon 9, and
- have submitted an abstract which has been accepted for oral or poster presentation.

Applicants who wish to attend the conference but will not be presenting a paper should explain how attending this conference relates to their academic goals. Selection will be on the basis of economic need and potential contribution to the conference.

Applicants should send an email letter of interest to the LabPhon 9 organizers at labphon9@uiuc.edu no later than March 1.

The LabPhon 9 organizers gratefully acknowledge support for these grants from the National Science Foundation.

University of Illinois at Urbana-Champaign | Department of Linguistics | Beckman Institute for Advanced Science & Technology

LabPhon 9
Change in Phonology

Ninth Conference on Laboratory Phonology
Thursday 24 June - Saturday 26 June, 2004

International Scholar Participation Grant

The organizers of the ***9th Conference on Laboratory Phonology (LabPhon 9)*** are pleased to announce the availability of one or two small grants to partially support the participation of international scholars from economically disadvantaged countries in LabPhon 9. These grants are intended to cover registration and (partial) travel expenses. To qualify, applicants must be affiliated with an institution from an academically underrepresented country and document insufficient support from their home institution. Preference will be given to applicants who submitted an abstract which has been accepted for oral or poster presentation at LabPhon 9. Applicants who will not be presenting a paper should justify the relevance of the conference to their professional development. Selection will be on the basis of economic need and potential contribution to the conference.

Applicants should send an email letter of interest to the LabPhon 9 organizers at labphon9@uiuc.edu no later than March 1.

The LabPhon 9 organizers gratefully acknowledge support for these grants from the National Science Foundation.

University of Illinois at Urbana-Champaign | Department of Linguistics | Beckman Institute for Advanced Science & Technology

WELCOME RECEPTION, WEDNESDAY JUNE 23RD

Conference participants are invited to a Welcome Reception on Wednesday, **June 23, 6-8 p.m**. at the **Beckman Institute West Atrium**. The updated conference program and other information will be available at the reception, and on-site registration will be processed for those who have not pre-registered. Appetizers and beverages will be served.

:: LabPhon 9 on the web ::
http://www.linguistics.uiuc.edu/labphon9

:: Contact ::
labphon9@uiuc.edu

Embracing Multilinguality: Defining Phonetic Features for Speech Technology

Julie Carson-Berndsen, Anja Geumann & Moritz Neugebauer
University College Dublin

In the past one of the major hurdles in trying to convince speech technologists to work with features rather than phones or phonemes (or even diphones or demi-syllables) as a unit of description has been that feature classifications have been either too abstract (i.e. not close enough to what actually occurs in the speech signal) or too phonetic (i.e. provide too much detail which cannot be integrated into a uniform model). While features have become more or less a standard way of describing sound characteristics in phonology and groundbreaking work done by Stevens [8] and many others has established features as referring to acoustic and/or articulatory characteristics, their use in speech technology applications has been avoided with relatively few exceptions [2,3,4,5,6,7]. While implicitly a lot of speech technology applications have made use of laboratory results, the interaction between implementation of results and their computational processing and screening has often proven to be difficult, if not impossible, due to nonconformity of structures.

We believe that the definition of a phonetically motivated feature set is the key to robust and scalable multilingual speech systems. We present an extensive set of features that covers languages as diverse as English, French, German, Irish, Romanian and Spanish, as well as to some extent allophonic variation within these languages. It replaces the current set of features used in our statistical feature extraction model [1] that is part of an agent-based recognition system [10]. The set of features aims to be sustainable for an implementation in a knowledge based system as well as for the use within a HMM based recognizer. Our main priority in the design of the multilingual feature set was to maximize scalability and reusability of the features in the development of different speech applications. Rather than taking a standard definition of sounds in terms of IPA features as found occasionally in the literature [9], we wish to make no *a priori* assumptions about the arity (e.g. unary, binary etc.) or clustering of the features into classes (e.g. manner, place etc.). Instead sounds are defined in terms of the presence of acoustic and articulatory phonetic properties and clustering of features and definition of feature implications are inferred automatically from the phonetic specifications allowing for generation of classifications, which exhibit significantly more structure than traditional feature classifications.

We acknowledge the fact that a rich set of features might not be easily accessible for manual optimization, such as identification of implicational relations between individual features. Therefore, we propose an algorithmic method based on automated deduction to compute these correspondences between individual features and furthermore between all sets of sounds created by combinations of those. We apply our algorithm to automatically generate feature hierarchies similar to type hierarchies in unification-based grammar formalisms, where features are ordered with respect to the size of their extension, i.e. the sounds they describe. While features that characterize large sets of sounds are assigned a rather high position in the hierarchy, features with a relatively small extension are situated at the bottom. We will show how this method will enable us to carry out consistency checking on our feature set since we automatically detect undefined features and redundant specifications. By these means, even a fairly large multilingual feature space can be maintained as well as it can be mined for language-dependent and language-specific phonological implications.

Finally, a structured multilingual feature set supports an organization of feature extractors which can maximize the information which can robustly be extracted from the speech signal. It will also supplement this information with inferences based on implications exhibited in the feature set. Our aim is that the multilingual feature space will form the basis for the design of novel feature extraction techniques tailored to the individual features.

References

[1] Abu-Amer, T. & J. Carson-Berndsen (2003): HARTFEX: A multi-dimensional system of HMM based recognizers for articulatory feature extraction. In: *Proceedings of the 8th International Conference on Speech Communication and Technology* (Eurospeech 2003), Geneva.

[2] Carson-Berndsen, J. (1998): *Time Map Phonology: Finite State Models and Event Logics in Speech Recognition*. Dordrecht, Holland: Kluwer Academic Publishers.

[3] Chang, S., S. Greenberg & M. Wester (2001): An elitist approach to articulatory-acoustic feature classification. *Proceedings of the 7th Eurospeech Conference on Speech Communication and Technology* (Eurospeech 2001), 1725-1728.

[4] Juneja, A. & C. Espy-Wilson (2003): An event-based acoustic-phonetic approach to speech segmentation and E-set recognition. In: *Proceedings of 15th International Congress of Phonetic Sciences*, Barcelona.

[5] Kirchhoff, K. (1999): *Robust Speech Recognition Using Articulatory Information*, PhD thesis, University of Bielefeld, Germany.

[6] Koreman, J., Andreeva, B. & Strik, H. (1999): Acoustic parameters versus phonetic features in ASR. In: *Proceedings of the 14th International Congress of Phonetic Sciences* (ICPhS'99), San Francisco.

[7] Lahiri, A. & H. Reetz (2002): Underspecified Recognition. In: C. Gussenhoven & N. Warner (eds.) *Laboratory Phonology 7*. Mouton de Gruyter: Berlin, New York, 637-676.

[8] Stevens, K.N. (1998): *Acoustic Phonetics*. MIT Press: Cambridge, London.

[9] Stüker, S., T. Schultz, F. Metze & A.Waibel (2003): Multilingual Articulatory Features. In: *Proceedings of* ICASSP 2003.

[10] Walsh, M. (2003): Recasting the Time Map Model as a Multi-Agent System. In: *Proceedings of 15th International Congress of Phonetic Sciences*, Barcelona.

A study of the melodic contour of Mandarin and French through Syntactic Structural Ambiguity

Ivan Chow
University of Toronto

An acoustic study was conducted to investigate the melodic properties of prosodic boundary markers in the resolution of Syntactic Structural Ambiguity (SSA) in Mandarin and French. Based on the experimental results, I present the following arguments: as a well-documented marker of discourse junctures in other languages [1], [6], [7], [17], [18], [21], pitch reset is also used within the sentence to convey prosodic boundaries in both Mandarin and French; the co-presence of language-specific melodic features and pitch reset (found in both languages) in the melodic contours and prosodic boundaries provides empirical support to the Overlay Model of Intonation (OMI) [8-10].

SSA refers to the phenomenon in which an utterance is represented by more than one well-formed syntactic structure cf. [11], [12], [19], [27]. In a situation where contextual information is not available for disambiguation, prosody is shown to be effective in resolving SSA [20]. The intended syntactic structure is conveyed by aligning the prosodic and syntactic boundary [14], [15], [20]. The acoustic experiments were conducted to examine the melodic devices used to convey major syntactic boundaries in speech production in Mandarin and French.

Mandarin: hong2se4 / zhi3 na2 / san1 zhang1 (Take only 3 sheets of the red ones.)
hong2se4 zhi3 / na2 san1 zhang1 (Take 3 sheets of red paper.)
French: Le petit / garde la porte. (The little one guards the door.)
Le petit garde / la porte. (The little guard wears it.)

A read-aloud corpus of carefully designed syntactically ambiguous sentences (such as the ones above) was collected. The sequences were designed so that the natural prosodic breaks align with major syntactic boundaries. In the script, contextual primer sentences precede the test-sequences in order to semantically and rhythmically prepare the readers to convey the intended structure. The purpose was to minimize unnatural *ad hoc* responses that would otherwise be elicited by explicitly marking the prosodic boundaries in the script.

Phonetic analyses of the corpus reveal that, although slightly less reliable than temporal cues, melodic cues are effectively used by speakers of both languages in marking prosodic breaks. Although the basic contour shape of Mandarin and French are governed by different language-specific phonological rules within different linguistic domains, pitch reset [8], [13], [16], [21] is proven to be a prominent prosodic boundary-marking device in both languages [4], [5]. As a well-documented marker of discourse juncture in narrative and conversational settings in many languages [2], [3], [17], [21-26], results indicate that pitch reset is present in junctures within the sentence in French and Mandarin as well.

The OMI states that the melodic contour is generated by a combination of two pitch components. In the *accent* component, pitch movement is largely governed by language-specific rules with rather 'localized' impact on smaller prosodic domains. The *phrase* component consists mainly of the language-universal declination contour and is subject to rules dominating larger domains with more widespread impact on melody. The observed language difference in melodic pattern and phonological rules governing the melodic contour coincide with the descriptions of the *accent* component. Pitch reset, an element of the language-universal *phrase* component, is

found to be a prominent boundary-marker in both languages. As a result, data from the present experiment provides further empirical support to the OMI.

In summary, the study of the melodic boundary markers of two typologically different languages showed pitch reset to be a prominent prosodic boundary marker within the sentence. The observed differences and similarities in terms of syntax-melody correspondence between Mandarin and French are shown to be in accordance with the view of the OMI.

Reference (Selected):

[1] Brown, G.; Currie, K.; Kenworthy, J., 1980. *Questions of Intonation*. Croom Helm, London.

[2] Brubaker, R. S., 1972. Rate and pause characteristics of oral reading. *Journal of Psycholinguistics* 1, 141-147.

[3] Bruce, G., 1982. Textual aspects of prosody in Swedish. *Phonetica 39*, 274-287.

[4] Chow, I., 2003. *Prosodic Structures in French and Mandarin*. Ph. D. Thesis. University of Toronto.

[5] Chow, I., to appear. Ambiguity resolution in French by way of melodic contour: An acoustic study. *Approaching Asymmetry in the Interface (AAI) Conference Proceedings*. Université du Québec à Montréal. Oct, 2003.

[6] 't Hart, J.; Collier, R.; Cohen, A., 1990. *A perceptual study of intonation: An experimental phonetic approach*. Cambridge University Press, Cambridge.

[7] Hirschberg, J.; Grosz, B., 1992. Intonational Features of Local and Global Discourse Structure. *Proceedings of the Speech and Natural Language Workshop*. DARPA, Harriman, NY. 441-446.

[8] Ladd, D.R., 1984. Declination: a review and some hypotheses. *Phonology Yearbook 1*, 53-74.

[9] Ladd, D.R., 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.

[10] Ladd, D.R., 2001. Intonation. Haspelmath, M., E. Konig, W. Oestereicher, W. Raible (eds.), *Language typology and language universals: An International Handbook 1-2*. Berlin, Germany: de Gruyter.

[11] Lehiste, I., 1973. Phonetic disambiguation of syntactic ambiguity. *Glossa 7*, 107-122.

[12] Lehiste, I., 1975. The phonetic structure of paragraphs. A. Cohen and S.G. Nooteboom (eds.). *Structure and process in speech perception*. Heidelberg: Springer. 195-203.

[13] Levelt, W. J. M., 1989. *Speaking : from intention to articulation*. Cambridge, MA. London : MIT Press

[14] Martin, Ph., 1981. Vers une théorie syntaxique de l'intonation. M. Rossi et al. (eds). *Intonation: de l'acoustique à la sémantique*. Klincksieck.

[15] Nicol, J.L., 1996. What can prosody tell a parser? *Journal of Psycholinguistic Research* 25:2, 179-193.

[16] Ohala, J. J., 1986. Phonological evidence for top-down processing in speech production. J. S. Perkell and D. H. Klatt (eds.), *Invariance and Variability in Speech Processes*. Lawrence Erlbaum Associates, Hillsdale, N.J. 386-401.

[17] Oliveira, M. Jr., 2003. Pitch reset as a cue for narrative segmentation. Amina

Mettouchi & Gaëlle Ferré, *Proccedings of IP2003, Prosodic Interfaces*. Nantes, France, 73-78.

[18] de Pijper, J. R.; Sanderman, A. A., 1994. On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues. *Journal of the Acoustical Society of America 96*, 2037-2047.

[19] Shen, X. S., 1992. A Pilot study on the relation between the temporal and syntactic structures in Mandarin. *Journal of the International Phonetic Association 22*:1/2, 35-43.

[20] Speer, S.R.; Kjelgaard, M.M.; Dobroth K. M., 1996. The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities. *Journal of Psycholinguistic Research 25*, 247-268.

[21] Swerts, M., 1997. Prosodic features at discourse boundaries of different strength. *Journal of the Acoustical Society of America 101:1*, 514-521.

[22] Swerts, M.; Collier, R.; Terken, J., 1994. Prosodic predictors of discourse finality in spontaneous monologues. *Speech Communication 15*, 79-90.

[23] Swerts, M.; Geluykens, R., 1994. Prosody as a marker of information flow in spoken discourse. *Language Speech 37*, 21-43.

[24] Terken, J., 1984. The distribution of pitch accents in instructions as a function of discourse structure. *Language Speech 27*, 269-289.

[25] Thorsen, N. G., 1985. Intonation and text in Standard Danish. *Journal of the Acoustical Society of America 77*, 1205-1216.

[26] Thorsen, N. G., 1986. Sentence intonation in textual context — supplementary data. *Journal of the Acoustical Society of America 80*, 1041-1047.

[27] Wenk, B.J. & Wioland, F., 1982. Is French really syllable-timed? *Journal of Phonetics 10*, 193-216.

On the emergence and loss of opacity effects in acquisition

Daniel A. Dinnsen
Indiana University

It has long been recognized that certain phonological processes behave asymmetrically with respect to context. That is, there are processes that can apply in phonologically or morphologically derived environments, but those same processes are blocked from applying in a phonologically identical nonderived environment (e.g., Kiparsky 1976). This asymmetry results in generalizations that are opaque, i.e., not surface-true. The characterization of opacity effects has, to varying degrees, challenged all contemporary theories of phonology, but none more than optimality theory (Prince & Smolensky 1993/2002). In an attempt to deal with these and other opacity effects, McCarthy (2002) put forward 'comparative markedness' as a revision to the theory. His proposal modifies original conceptions of markedness by splitting each markedness constraint into two independent and freely permutable constraints that assign violation marks differentially based on a comparison with the fully faithful candidate. While opacity effects are relatively common in fully developed languages and while comparative markedness enjoys some success in accounting for them, surprisingly little is known about the nature of these effects in acquisition. Do comparable opacity effects occur in the early stages of acquisition? If so, how and why do they emerge, and how are they lost? How does comparative markedness fare in accounting for these acquisition facts?

This paper addresses these questions by focusing on longitudinal data from a child who exhibited the emergence and loss of an opacity effect that was not present in the primary linguistic data of the target language. The child (age 4;4) was normally developing in all respects, except for evidence of a phonological delay. The child presented with two commonly occurring and interacting error patterns: Word-initial affricates deaffricated (Deaffrication), and word-initial coronal stops assimilated in place to a following (nonadjacent) dorsal consonant (Place Assimilation). The interaction of these error patterns at the presenting stage resulted in perfectly transparent outputs (i.e., word-initial affricates deaffricated and, along with coronal stops, assimilated when followed by a dorsal). The child was enrolled in a clinical treatment study employing a single-subject experimental design that was aimed at eradicating Place Assimilation. As expected, Place Assimilation was largely eliminated following treatment, but Deaffrication continued. Surprisingly, however, Place Assimilation persisted only for those stops that were derived from Deaffrication; nonderived stops were immune to Place Assimilation. This is an example of an opacity effect which McCarthy (2002) has dubbed a 'grandfather effect'. Subsequently and without further treatment, Deaffrication and the opaque remnants of Place Assimilation were lost.

The claims of comparative markedness are evaluated against these and other acquisition facts and are shown to account for the emergence and loss of grandfather effects in a straightforward fashion. The finding that grandfather effects emerge in early stages of acquisition despite the absence of those effects in the target language is important in itself because it bears directly on the long-running debate over the role of the input. As a result of these considerations, some new insight is also gained about the emergence of another commonly occurring (but theoretically problematic) class of opacity effects in acquisition, namely chain

shifts (e.g., interdentals are realized as [f], but /s,z/ are realized as interdentals). Importantly, it is the derived sound in these cases that is blocked from undergoing a process. The suggestion is that, while a chain shift may not occur in the input, it emerges naturally as a resolution to conflicting universal constraints.

Selected References

Kiparsky, Paul. 1976. Abstractness, opacity and global rules. The application and ordering of grammatical rules, ed. by A. Koutsoudas, 160-86. The Hague, Netherlands: Mouton.

McCarthy, John J. 2002. Comparative markedness. Ms., University of Massachusetts, Amherst [Rutgers Optimality Archive #489].

Prince, Alan and Smolensky, Paul. 1993/2002. Optimality Theory: Constraint interaction in generative grammar. Report No. RuCCS-TR-2. New Brunswick, NJ: Rutgers University Center for Cognitive Science.

Something from Nothing: Permutability of Competence and Performance in Phonological Acquisition

Judith A. Gierut
Indiana University

A long-held view is that perception precedes production in language acquisition. The reverse relationship does occur, however, at least as reported for second language acquisition (Gass, 1984). Interestingly, there is new evidence in first language acquisition to suggest that typically developing children produce semantic and syntactic structures in the absence of corresponding knowledge (Coles-White, deVilliers, & Roeper, 2003). The purpose of this paper is to report comparable effects within the domain of phonology, whereby children's knowledge of phonemic contrasts lags productive use of those same contrasts. The phonological property of interest is the phonemic distinction between /s/ and theta in English; namely, the stridency contrast among coronal continuants. Our goal was to determine what preliterate preschoolers know about this later acquired distinction, and how this knowledge maps onto their productive phonology. Forty typically developing children were recruited: 20 maintained the s-theta contrast productively as determined by extensive spontaneous speech samples and structured probes, and 20 others of the same age did not (*M* CA 4;3). Twenty-three adult speakers of English also participated (*M* CA 22;3). Three experiments were conducted using a conventional constrained classification task (cf. Treiman & Breaux, 1982). Participants were presented with triplets of stimuli and asked to select which 2 of 3 were similar; perceived similarity was thus the dependent variable. Adults responded in paper/pencil format, and children played a triad game, whereby three identically pictured characters were mounted in corners of a triangular board. They were to listen to each character's 'name' and then select which two were 'friends.' Character 'names' were CV syllables that systematically varied in consonantal place of articulation, continuancy and stridency as the independent variables. Three experiments were conducted to establish the relative ranking of these featural dimensions by each group of participants. Across experiments, 3 of 4 logically possible relationships between perceptual similarity and production were instantiated. Specifically, children who did not produce the s-theta contrast did not utilize the stridency distinction in their judgments of perceptual similarity (–perception/–production). Moreover, the same was true for children who used the s-theta contrast productively. This was the crucial observation: stridency was not relevant to children's similarity judgments, even though it was central to maintaining distinctions of their productive phonological systems (–perception/+production). Finally, as expected, adults relied on stridency as a pertinent dimension in both their similarity judgments and production (+perception/+production). These results hint at developmental changes in the relationship between perception and production that are associated with the composition of the phonological system, as well as chronological age. They also raise questions about the kinds of mechanisms that a theory of phonological development must employ to yield such patterns. Consistent with the current framework of optimality theory, emergence of the unmarked will be considered as a possible solution with implications for claims about the fixed ranking of constraints.

References

Coles-White, D., deVilliers, J., & Roeper T. (2003, November). *Emergence of barriers to wh-movement, negative concord and quantification*. Paper presented at the BU Conference on Language Development, Boston, MA.

Gass, S. (1984). Development of speech perception and speech production abilities in adult second language learners. *Applied Psycholinguistics*, 5, 51-74.

Treiman, R., & Breaux, M. (1982). Common phoneme and overall similarity relations among spoken syllables: Their use by children and adults. *J. Psycholinguistic Research*, 11, 569-598.

Perceived similarity: universal or language-dependent?

Christine Haunz
University of Edinburgh

Listeners' perceived similarity may be influenced by different factors, which fall into two major groups; the first group are universal factors, which correlate with acoustic-phonetic characteristics, such as the availability of cues signalling sound contrasts, in specific environments. For example, cues to the place of articulation (POA) of plosives are clearer when the plosive is released into a following vowel than when it is unreleased, as it may be word-finally. Thus a POA difference is predicted to be more similar word-finally than word-initially. These factors are assumed to be true universally, and thus to influence judgments of similarity cross-linguistically.

The second group of factors takes into account the language-specific organization of the phoneme inventory and phonotactics. The perceptual difficulties that language learners have with foreign phonemes have been shown by a wide range of studies on second language acquisition (see Strange 1995). An influence of native phonotactics is contested; Silverman (1992) admits only segmental influence, whereas Dupoux et al. (1998) claim that listeners perceive foreign words according to their own phonotactic rules, e.g. in the case of Japanese with vowels intervening between adjacent consonants.

The latter view predicts that non-native speakers rate the similarity between forms that are illegal in their native language (L1) and those that they perceptually assimilate them to as significantly higher than listeners of a language in which both forms are legal. Conversely, Steriade (2001), Hayes and Steriade (2003) and Jun (1995) predict universal hierarchies of the relative perceptibility of certain contrasts (in specific environments). To determine the relative influence of universal and specific factors, an experiment comparing the perceived similarity of same sound pairs by speakers of different languages was conducted. Both types of predictions were tested in a cross-linguistic experiment, in which 8 native speakers each of English and Russian rated the similarity of pairs of Russian pseudo-words differing in the onset in the following ways:

1) C_1C_2 vs. C_1VC_2 (where the epenthesised V was a schwa)
2) C_1C_2 vs. $C_{1/2}$
3) C_1C_2 vs. C_1C_3 / C_4C_2 (only one feature change per comparison)
4) C_1VC_2 vs. C_1VC_3 / C_4VC_2

Results showed evidence in favour of both sides. There was no significant difference between English and Russian speakers when one of the pair was illegal and one legal in English. On the other hand, a difference was found in the case of both words containing illegal clusters; here the similarity ratings of English listeners were significantly higher than those of the Russian listeners. This suggests that phonotactic perceptual assimilation does not necessarily occur, but in cases where at least one stimulus can be mapped onto a native form, the perception of differences is not impaired. This supports the claim that language background has a bearing on the way listeners perceive differences and similarities.

Evidence supporting universal perceptibility hierarchies was also found: a voicing change is less perceptible than a change in nasality for the English as well as the Russian group. POA changes are more perceptible before coronals than non-coronals, and for fricatives as compared to plosives, irrespective of language. The deletion of a consonant is a more severe alteration than

the addition of a vowel (schwa). Within the group of consonant deletions, the deletion of the prevocalic consonant (C_2) was more perceptible than that of C_1. However, a comparison of the ratings of cluster pairs and the corresponding epenthesised onset pairs (3 vs. 4) only showed the predicted higher perceptibility for the case of two released consonants (4) for the Russian, but not the English listeners. These findings suggest that there is a division of labour between the universal acoustic-phonetic and the language-specific factors, and that they interact to determine perceived similarity.

References

Dupoux, Emmanuel, Kakehi, Kazuhiko, Hirose, Yuki, Pallier, Christophe, & Mehler, Jacques (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, 25(6), 1568-1578

Hayes, Bruce & Steriade, Donca (2003, in press). The phonetic bases of phonological markedness, in Bruce Hayes, Robert Kirchner and Donca Steriade (eds.), *Phonetically Based Phonology*, Cambridge University Press.

Jun, Jongho (1995). Perceptual and articulatory factors in place assimilation: an optimality-theoretic account. PhD dissertation, UCLA. [http://yu.ac.kr/~jhjun/]

Silverman, Daniel (1992). Multiple scansions in loanword phonology: evidence from Cantonese, *Phonology 9*, 289-328

Steriade, Donca (2001). The Phonology of Perceptibility Effects: the P-map and its consequences for constraint organization. [http://www.linguistics.ucla.edu/people/steriade/papers/P-map_for_phonology.doc]

Strange, Winifred (1995). Cross-language studies of speech perception: A historical review, in: Strange, Winifred (ed.) *Speech perception and linguistic experience: Issues in cross-language research*, Baltimore: York, 3-45

Prosodic and phonotactic influences on fricative voicing assimilation in German

Claudia Kuzla & Taehong Cho
Max Planck Institute for Psycholinguistics

Recent phonetic studies have shown that the phonetic realization of segments varies with the prosodic position in which they occur. For example, domain-final segments are longer than domain-initial ones; segments are produced more strongly domain-initially than medially; these effects increase in magnitude as the prosodic domain becomes larger, i.e. phrase-initial segments are stronger than word-initial (but phrase-medial) ones. (Edwards, Beckman & Fletcher, 1991, JASA; Fougeron & Keating, 1997, JASA; Cho & Keating, 2001, JPhon; Cho, 2003, 15th ICPhS).

In this study, we examined the acoustic realization of the initial German fricatives /f, v, z/ in different prosodic positions. In particular, we investigated whether the assimilatory devoicing of voiced fricatives /v, z/ after a voiceless obstruent (e.g., /z/ -> /[s]/ t_) is a gradient process sensitive to prosodic structure, i.e. varying with the size of intervening prosodic boundaries. It was hypothesized that a smaller prosodic boundary between segments increases the cohesion between them such that the degree of voicing assimilation of /z, v/to voiceless /t/ is greater across a smaller prosodic boundary. In addition, we examined how the hypothesized variation in voicing assimilation due to prosodic position is further conditioned by a phonotactic constraint: In German, while both voiced and voiceless labiodental fricatives /v, f/ occur word-initially, only the voiced alveolar fricative /z/ (not /s/) is allowed in that position. Excessive assimilatory devoicing of voiced /v/ would confound the phonological contrast with its voiceless counterpart /f/, which would constrain the voicing assimilation of /v/. In contrast, no such restriction would apply to /z/, allowing for more voicing assimilation.

Ten speakers read various sentence types intended to elicit different prosodic boundaries. Two versions of each sentence type were used: one containing a stop-fricative sequence (e.g., /…t#v…/, # = prosodic boundaries, *ha*[t]#[v]*?lder* 'has forests'), the assimilation environment, and one containing a vowel-fricative sequence (e.g., /…E#v…/, *hatte*#[v]*?lder* 'had forests), as a control condition. The elicited utterances were grouped into three prosodic categories (Major, Minor, and Word), with regard to the size of the prosodic boundary in the test sequences. The Major boundary was defined as having both a pause and an intonational marker ("boundary tone"); the Minor boundary as having no pause, but an intonational marker; and the Word boundary as having neither a pause nor an intonational marker. Two speakers were excluded from the analysis because they did not produce any tokens belonging to the Minor boundary. Acoustic measurements included preboundary syllable duration, fricative duration, and the amount of voicing as measured by percent (%) of voicing during the fricative.

Results are as follows. First, the three-way prosodic grouping was supported by the preboundary (final) lengthening pattern of Major > Minor > Word. Second, there was a main effect of boundary size on the fricative duration, but the effect was not completely in accordance with the domain-initial strengthening patterns, showing only a pattern of Minor . Word. Third, the relative amount of voicing showed a voicing assimilation effect: both /v/ and /z/ were realized with less voicing after /t/ than after a vowel. Considering /z/, the amount of voicing assimilation varied with prosodic boundary size, showing a general pattern of Major > Minor > Word. This supports the hypothesis of voicing assimilation as a gradient process sensitive to prosodic structure. However, when /v/ was considered, no such effect was found. This asymmetry between /z/ and /v/ suggests that the effect of prosodic boundary is constrained by the

language-specific phonotactics, such that when there is a need to maintain a phonological contrast (/v/ vs. /f/), the boundary effect is suppressed.

The results of this study suggest that phonetic and phonological processes, as found in the fricative duration and voicing assimilation in German, are bounded by the interplay between prosodic structure and phonotactics of the language.

Two complementary phonotactic constraints interfacing grammar with phonetics

Eleonora Cavalcante Albano
Universidade Estadual de Campinas

In Southeastern Brazilian Portuguese (henceforth SBP), the distribution of the open mid vowels [E, O], which occur only under stress, automatically increases when regular verbs are formed from stems ending in [e, o], due to a verb paradigm alternation well known to generative phonology (e.g., Harris 1974, Maia 1981). In addition, at least two other productive processes create new open mid vowels: acronym pronunciation (e.g., 'M[É]C', Ministry of Education), and noun or adjective derivation via inherently unstressed stems or suffixes (e.g., 'son[Ó]grafo', sonagraph).

This paper proposes a unified account of all of these facts. Two rules of the earlier studies – vowel opening and close vowel harmony – are recast as probabilistic phonotactic constraints (Pierrehumbert 2003) differentially affecting grammatical classes and interfacing grammar with the phonetics of prosody.

The data come from two electronic corpora of SBP. The first is a word frequency count available on the internet (Linguateca, n. d.), which compiles 2 million words from a major Southeast Brazil newspaper. This was filtered to remove spurious biases and passed through a very accurate orthography-to-phone converter. The second is a treatment of Ferreira's (1977) 27,000 word dictionary that incorporates both phonetic script and grammatical class coding.

Words with at least one vowel preceding lexical stress were selected from both corpora. Positional probabilities were estimated for rightmost prestressed /i, e, a, o, u/ and stressed /i, e, E, a, O, o, u/. Both corpora indicate that /E, O/ are around 60% less frequent than /e, o/ under stress, a disadvantage which apparently runs counter the rate at which their distribution is growing. Yet, if open vowels are pooled together, as suggested by the alternation pattern, the stressed vowel distribution is significantly ($p<.0001$) skewed towards openness.

The issue was further pursued by looking into the vowel phonotactics of the grammatical subsets in the dictionary. V-to-'V observed and expected probabilities were estimated and compared for the total set as well as for the noun, adjective, and verb subsets.

Segmental analysis of the V-'V pairs, achieved via statistical treatment of observed frequencies and O/E ratios, reveals that grammatical classes have different biases. Analogous subsegmental analysis of the pairs, achieved after coding them along the dimensions of opening, place, and rounding, further reveals that the total corpus shows no significant bias for vowel harmony or disharmony on any dimension. By contrast, verbs are significantly biased for close vowel harmony ($p<.001$) while adjectives ($p<.001$) and nouns ($p<.05$) are biased for opening disharmony. An ensuing study of productive suffixes further showed that disharmony towards openness is also preferred by derivation in all three classes ($p<.0001$).

Productive derivation thus conspires with major regular inflection in /a/ to create open vowels in stressed position. In turn, stressed /e, o/ in verbs mark minor regular inflection in /e, i/, and conform to an iterative right-to-left trend for stem vowels to agree in opening with such markers ($p<.001$ for V1 and $p<0.05$ for V2), accounting for the verb close vowel harmony bias. It is as if the mid vowels of the verb paradigm resulted from a trade-off between the grammar and the phonetics. The infixation of 'close' in a mid vowel occurring rightmost in a stem where /e/ or /i/ is missing guarantees inflectional transparence while enhancing derivational opaqueness in the minor classes. In contradistinction, the open vowels found elsewhere in the verb paradigm

enhance articulatory economy and perceptual salience (in the sense of Recasens 2003) in the major class. This interpretation rests on two well documented phonetic facts: (1) the universal positive correlation between vowel opening and duration holds for SBP (Barbosa 2002); (2) there is a growing body of evidence showing that the only stable acoustic correlate of stress in SBP is a linear increase in normalized V-to-V duration toward the stressed vowel (Barbosa & Arantes 2003).

References

Barbosa, P. A., 2002. Explaining cross-linguistic rhythmic variability via a coupled-oscillator model of rhythm production. Proceedings of the Speech Prosody 2002 Conference. Aix-en-Provence, France, pp. 163-166.

Barbosa, P. & P. Arantes. 2003. Investigation of Non-Pitch-Accented Phrases in Brazilian Portuguese: No Evidence Favoring Stress Shift. In: Solé, M. J.; D. Recasens & J. Romero (eds.). Proceedings of the XVth International Congress of Phonetic Sciences. Barcelona: Universitat Autònoma de Barcelona, pp. 475-478.

Ferreira, A. 1977. Minidicionário Aurélio. Rio de Janeiro: Nova Fronteira.

Harris, J. 1974. Evidence from Portuguese for the 'Elsewhere Condition' in phonology. Linguistic Inquiry, 5(1): 61-80.

Linguateca, n. d. http://acdc.linguateca.pt/cetenfolha/

Maia, E. 1981. Phonological and lexical processes in a generative grammar of Portuguese. Brown University unpublished doctoral dissertation.

Pierrehumbert, J. 2003. Probabilistic Phonology: Discrimation and Robustness. In: Bod, R.; J. Hay & S. Jannedy (eds.). Probability Theory in Linguistics. The MIT Press, pp. 177-228.

Recasens, D. 2003. Articulation and Sound Change in Romance. In: Solé, M. J.; D. Recasens & J. Romero (eds.). Proceedings of the XVth International Congress of Phonetic Sciences. Barcelona: Universitat Autònoma de Barcelona, pp. 231-234.

Sixty years of bilingualism affect the pronunciation of Latvian vowels

Z.S. Bond*, Dace Markus† & Verna Stockmal*
Ohio University and University of Latvia†*

For fifty years after WW II, Latvia was incorporated into the former Soviet Union. Although in theory the use of regional languages was not discouraged, in practice knowledge of Russian was obligatory. It was nearly impossible to obtain government services, get medical care, or even make routine purchases without knowing Russian. Since 1991, Latvian has again become the official language and knowledge of Russian is widespread but optional.

These political events have created a natural experiment in the effects of almost universal bilingualism on a language. To assess the impact on pronunciation, ten native speakers of Latvian, long-term residents of Riga, were recorded reading a word list and a short narrative, designed to inventory the phonological patterns of Latvian. The recordings were made in 2001, ten years after the collapse of the Soviet Union. Two talkers represented each generation since the end of WW II. The oldest talkers were of retirement age; they would have been young adults at the time when Russian became an obligatory language and provide a baseline for the pronunciation of Latvian before extensive bilingualism became the rule. The intervening three groups were professionals, who received their education and established themselves in their fields during the time of maximum pressure from Russian. The youngest talkers, high school students, have grown up while Russian is no longer the politically dominant language.

Vowels differed across the generations, both in duration and formant structure. All talkers employed contrastive long and short vowels, but the ratio changed from almost 2.5 for the oldest to 1.5 for the youngest, suggesting that the younger talkers are not maintaining this contrast as clearly as the older talkers do. All talkers also employed the vowels of the Latvian inventory in stressed and unstressed syllables but the younger talkers produced vowels with a higher F2 for back vowels and a lower F2 for some front vowels indicating a tendency towards vowel reduction. The middle generation exaggerated the contrast between /e/ and /ae/, as if to differentiate their pronunciation of Latvian from Russian.

Instruction and Phonetic Change

Josefina Carrera-Sabaté
Universitat de Barcelona and Universitat de Lleida

Catalan is a Romance language spoken, among other places, in Catalonia, in the north-east of the Iberian Peninsula. From the 1940s till nowadays, Catalan has changed its status, from a language prohibited in public affairs but spoken in interpersonal relationships, to becoming a language taught in schools and reaching the status of schooling language. This historic and social situation has resulted into Catalan-speaking people who can not write Catalan, Catalan-speaking people who have learnt written Catalan for work or personal reasons and, finally, Catalan-speaking people who took schooling in Catalan and use this language for all purposes. This state of affairs has made the Catalan society enhance the written language and, as a consequence, the presence of clear influences of the written language onto the spoken language is easily perceptible. (See Segarra, 1985; Romero, 1995; Carrera-Sabaté, 2002).

Taking into account this sociolinguistic situation, in this paper I present to what extent the instruction factor, linked to several social variables, is responsible for a variation found in *lleidatà*, a Catalan dialect spoken in western Catalonia. The process of phonetic variation takes place in the 3rd person singular morpheme in verb forms of the indicative imperfect tense, in the conditional mode and also in the indicative present tense in verbs of the first conjugation. Traditionally, in *lleidatà* 1st and 3rd person singular morphemes were distinguished by means of the following vowels: open *e* (1st person) and close *e* (3rd person). However, nowadays these endings tend to lose its distinctive value and to be emitted with only one final form: open *e*. The relevance of this incipient variation can be explained through the influence of the written language, since, in *lleidatà*, most unstressed <a> are pronounced as open *e*. (See Pueyo, 1976; Veny, 1982; Turull, 1990).

The study's starting point is 30 speakers from Alguaire, a village located 15 km away from western Catalonia's capital, Lleida. The social variables that I took into account in this stratified sample are: age, qualifications, knowledge of Catalan, gender and sociocultural status. The speakers' ages range from 4 to 70 years old (i.e. born between 1999 and 1933) and particular relevance was given to the first generations to take schooling in Catalan. The informants' productive and perceptual competence were quantified by means of the Goldvarb 2001 program following these steps:

a) On the one hand, I analysed the responses elicited out of a corpus of questions on several phenomena variable in *lleidatà*, among which there were 12 verbs –representing all the conjugations– that were to be uttered in indicative present, indicative imperfect and conditional. On the other hand, a total of 15 hours of more informal speech was recorded, this being obtained out of two protocols: a labovian style semi-directed interview and the interview that was generated during the meeting with each informant, following the proposals by Briggs (1986) and Hazen (2000).
b) The perception survey enabled an analysis of, in the first place, the degree of normality that the informants attributed to the open *e* solution once it was emitted in four words from the following sentences: *Creia que estudiaria molt* ('S/He though s/he would study very much') and *No troba mai el que li agrada* ('S/He never finds what s/he likes'). I focused on whether this vowel was considered either negatively or as a common form. Secondly, the audition of the verb forms *ella perdia* ('she lost') and *ella estudiava* ('she studied') emitted twice

consecutively –first with a final close *e* and then with an open one– made me realise: 1) the informants' capability to discriminate phonetic differences and 2) the evaluation carried out by the speakers about the vowels heard.

The results show the existence of a tendency to emit and accept the written-language-like forms by younger informants, the ones having schooling in Catalan, and also by those who had knowledge of written Catalan.

REFERENCES

BRIGGS, C. (1986) Learning How to Ask: A Sociolinguistic Apprasial of the Role of the Interview in Social Sicence Research. Cambridge University Press. Cambridge.

CARRERA-SABATÉ, J. (2002) Escola catalana i variació fonètica. Una evolució del vocalisme àton a Alguaire i a Lleida. Pagès Editors. Lleida.

HAZEN, K. (2000) Identity and Ethnicity in the Rural South: A Sociolinguistic View through Past and Present be. Publication of the American Dialect Society 83. Duke University Press. Durham.

PUEYO, M. (1976) L'estructura fonològica del vocalisme lleidatà. (inedit work. Lleida).

ROMERO, S. (1995) Escolarització i variació intralingüística: el cas de l'article etimològic en el català nord-occidental. Jornades sobre llengua i ensenyament. I: 225-236.

SEGARRA, M. (1985) Història de l'ortografia catalana. Empúries. Barcelona.

TURULL, A. (1990) Notes sobre el lleidatà. Ilerda Humanitats. XLVIII: 201-206.

VENY, J. (1982) Els parlars catalans. Raixa. Palma de Mallorca.

Sound change as a result of mimetic interactions between speakers: Implementation of imitation in the laboratory

Véronique Delvaux[†‡], Didier Demolin[‡*] & Alain Soquet[‡]
Université de Mons-Hainaut[†], Universidade de Sao Paolo and Université Libre de Bruxelles[‡]*

The general framework of this research is the emergence and the collective structuring of phonological systems. In this study, we try to identify some of the mechanisms involved in the propagation of sound changes within a speech community. Our main hypothesis is that sound systems result from mimetic interactions between individuals. Mimesis is defined as a supramodal, motor-modelling skill which creates representations that are retrievable from memory. Its function is to represent events in a conscious, intentional and deliberate way (Donald, 1991). In the context of phonology, mimesis is considered as the competence to develop and amplify variation.

Soquet et al. (2003) developed a methodology to generate and measure the modifications of phonetic realizations they recorded from speakers exposed to a simple interactive situation of communication. This study elaborates on these preliminary results. The objective is to characterize the dynamics of the speakers imitation behavior.

The experiment involves speech productions coming from two dialectal groups of the French-speaking Belgian community, further referred to as the "source" dialect and the "target" dialect. Two speakers of the target dialect are first recorded : R1, R2. Eight speakers of the source dialect are selected for the experiment : S1 to S8.

The experimental setting is the following: ideograms are presented on a computer screen. The computer randomly selects one participant to name an ideogram (within a carrier sentence). A participant can either be one representative of the source dialect (Sn), or a prerecorded speaker of the target dialect (R1 or R2).

Each speaker Sn of the source dialect participates in three successive phases: (1) Sn performs the task on his own; (2) the ideograms are named alternatively by Sn and by prerecorded R1 and R2; (3) Sn performs the task alone again. Speakers are told they take part in a memory/attention task.

The data are processed according to the method proposed in Soquet et al. (2003). The recordings of phase (1) are segmented manually; this segmentation is used to segment automatically the remaining items with dynamic programming. Several segment-related acoustic cues are extracted: segment duration, F0 contours, MFCC. Discriminant analysis is carried out based on the recordings of phase (1) in order to define the space that discriminates best between the items from the source dialect and the target dialect. Then, recordings of phases (1), (2) and (3) are compared within this discriminant space. The significance of the variance across phases is studied using Anova.

This design allows to record the usual phonetic realizations of an individual (phase 1), then to quantify their potential evolution when the subject is exposed to a dialect different from his own (phase 2), and finally to assess the extent to which the potential modifications are preserved even in the absence of the stimulus, i.e. of productions from the target dialect (phase 3). Shortly speaking, phase (2) and phase (3) respectively address the issues of imitation and mimesis.

References

Soquet A. , Demolin D., Delvaux V.(2003). "Mimetic interactions between speakers : an experimental study". 15th International Conference in Phonetic Sciences Book of Abstracts, Universitat Autonoma de Barcelona.

Donald M. (1991). Origins of the Modern Mind: three stages in the evolution of culture and cognition, Harvard University Press, Cambridge.

Accent Changes in an Endangered Japanese Dialect

Haruo Kubozono*, Nobuko Kibe† & Ichiro Ota†
Kobe University and Kagoshima University†*

This is a report on ongoing accent changes in Kagoshima Japanese, an endangered dialect spoken in the southern edge of Japan. This dialect has a very unique prosodic system among many Japanese dialects: e.g., it is one of the very few 'syllable' dialects of Japanese whose prosodic patterns rely entirely on the syllable, not the mora. Major characteristics are summarized in the table below, where 'Tokyo' represents most dialects of Japanese (Kibe 1997, 2000; Kubozono 1999, 2003).

(1) Prosodic Feature	**Kagoshima**	Tokyo
a. Basic unit	**Syllable**	Mora
b. No. of contrastive tonal patterns	**Two**	Multiple
c. Key element determining compound accent pattern	**First element**	Final element
d. Domain of accent/tone assignment	**Phrase**	Word

Partly because of these unique prosodic features, Kagoshima Japanese had resisted until recently the otherwise strong influence of Standard Tokyo Japanese. However, this is no longer true with its speakers of a relatively young generation who now use the vocabulary of Standard Japanese. To see how precisely its unique prosodic system is developing under the strong influence of Tokyo Japanese, we conducted fieldwork in five major points of the dialect-speaking area by having 40 plus speakers (20 speakers of 13 or 14-years old, 10 speakers in their 40s and early 50s, and 10 speakers in their 70s and 80s) in each point pronounce more than 500 test words (simplex words/compounds/phrases).

Analysis of the accent patterns produced by a total of 200 plus speakers has yielded several interesting results. First, the basic features in (1a) and (1b) are preserved largely intact, which suggests that the core structure of a dialect's prosodic system tends to be exempt from a direct influence of a socially dominant dialect. On the other hand, the most affected aspect of the accent system in the youngest generation concerns the matching of the two accent patterns with individual words. Many words traditionally belonging to Pattern A (high tone on the penultimate syllable, followed by a pitch fall) are pronounced with B Pattern (high tone on the final syllable, with no pitch fall), and vice versa. A closer examination has revealed that the changes are not arbitrary but exhibit an extremely high degree of correlation with tonal distinctions in Tokyo Japanese: The accent change from Pattern A to Pattern B is characteristic of words which are unaccented (or words involving no sudden pitch drop) in Tokyo Japanese, whereas the accent change from Pattern B to Pattern A occurs predominantly in words that are accented (i.e. pronounced with a sudden pitch drop) in Tokyo. Interestingly, a similar bi-directional change is observed in the pronunciation of compound nouns, thus breaching the rule in (1c) in part. The two types of changes are exemplified below in comparison with the corresponding Tokyo patterns: syllable boundaries are indicated by dots and high-pitched portions are capitalized.

(2) a. bu.ra.ZI.ru ? bu.ra.zi.RU 'Brazil' (Tokyo: bu.RA.ZI.RU)
sya.KAI.too ? sya.kai.TOO 'Socialist Party' (Toyko: sya.KAI.TOO)

b. doo.na.TU ? doo.NA.tu 'donut' (Tokyo: DOo.na.tu)

a.o.sin.GOO ? a.o.SIN.goo 'green light' (Tokyo: a.O.SIn.goo)

This suggests that a force for perceptual similarities underlies the initial prosodic changes that a socially less dominant dialect undergoes under the influence of a more dominant dialect and that these perceptually-driven changes occur within the prosodic framework of the original dialect, instead of borrowing the prosodic patterns of the dominant dialect in a direct manner.

Selected References

Kibe, Nobuko (ed.) (1997) Kagoshima-ken no Kotoba (The Language of Kagoshima Prefecture). Tokyo: Meiji Shoin.

Kibe, Nobuko (2000) Seinanbu Kyuushuu Nikei Akusento no Kenkyuu (Studies on the two-class-accent dialects of Southwest Kyuushuu). Tokyo: Benseisha.

Kubozono, Haruo (1999) Nihongo no Onsei (The Sound System of Japanese). Tokyo: Iwanami.

Kubozono, Haruo (2003) What does Kagoshima Japanese tell us about Japanese syllables ? A Festschrift for Dr. Masayoshi Shibatani on the Occasion of his 60th Birthday. Tokyo : Kurosio.

Phonetic variation in Missouri: The low vowel merger in transition

Tivoli Majors
University of Missouri

Across the western United States, the phonemic distinction between the vowels /a/ as in *cot* and /open o/ as in *caught* is being lost in a sound change called the low vowel merger. This merger is actively spreading in the state of Missouri. Speakers in the St. Louis metro area and the area immediately to its west are resisting this merger, while in the rest of the state the merger is at or nearing completion (Gordon 2001).

In this study, I have conducted acoustic analysis of word list data from 12 adolescent speakers from around the state: 3 unmerged speakers in the St. Louis metro area, 8 merged or nearly merged speakers from areas well outside of St. Louis, and one speaker (MS) living just outside of St. Louis in what I'll call a transition zone. Each speaker read a word list of 140 tokens; 48 words containing one of the relevant low vowels were analyzed. The following acoustic measures were examined at three time points (onset, midpoint, and offset): F1, F2, F3, as well as F0. Duration was measured as well. In order to strictly control for phonetic environment, statistical comparison was limited to 4 minimal pairs such as *Don-dawn* and 4 near minimal pairs such as *taller-dollar*.

Results of paired t-tests indicate that the speakers in the St. Louis metro area distinguish these vowels on the basis of place of articulation as measured by F1, F2, and F3. No differences were found in F0 or duration. For the most part, speakers well outside of the St. Louis area show absolutely no distinctions in any of the acoustic measures examined, indicative of a merger of the two phonemes. The analysis of the data for speaker MS in the transition zone was more complicated.

Because MS is found in an area of Missouri sandwiched between merged and unmerged speakers, data from her reading passages (RP), word list (WL) and minimal pair tasks (MP) were examined. Results indicate a very interesting shift in her vowel system. Analysis of the RP data indicates no differences between /a/ and /open o/ in this speech context. Analysis of the WL data, on the other hand indicates that she makes a distinction between these two vowels by relying solely on duration rather than on formant values as her St. Louis counterparts do. Auditory inspection of minimal pairs such as *cot-caught* indicates a salient difference in vowel length. The mean difference between the two vowel categories was 35 msecs, well beyond the just noticeable difference of 25 msecs established by Klatt (1976). In contrast, in the MP context, formant values differed significantly for the two vowels but duration did not. It would appear than in the more casual RP context, MS exhibits a merged system. In the more monitored MP speech situation, her system completely matches that of her St. Louis counterparts. However, in the intermediate WL speech context, her vowel system lies somewhere in between, signaling vowel identity only with durational cues, which have been argued to be less robust cues to vowel identification than formant frequencies (eg. Zahorian et al. 1999).

Others have been reported to switch between merged and unmerged low vowels (eg. Labov et al. 1972, Di Paolo 1992), but the direction of the shift for MS is opposite that found in the other studies. MS switches from a merged system in more casual settings (reading passages) to an unmerged system in more formal settings, indicating a prestige associated with the unmerged system of St. Louis. Her reliance on durational cues to signal vowel identity in one

speech context and formant cues in another indicates a degree of control and subconscious awareness of the low vowel merger not before reported.

References

DiPaolo, Marianna. 1992. Hypercorrection in response to the apparent merger of (open o) and (a) in Utah English. Language and Communication, 12, 3/4, p. 267-292.

Gordon, Matt. 2001. Missouri is for Mergers. Paper presented at NWAVE 30, North Carolina State University.

Klatt, Dennis H. 1976. Linguistic uses of segmental duration in English: acoustic and perceptual evidence. Journal of the Acoustical Society of America 59, p. 1208-1221.

Labov, W., M. Yaeger, and R. Steiner. 1972. A quantitative study of sound change in progress. Report on National Science Foundation Contract GS-3287. US Regional Survey, Philadelphia.

Zahorian, Stephen A. and Amir Jalali Jagharghi.1999. Spectral-shape versus formants as acoustic correlates for vowels. Journal of the Acoustical Society of America 94 (4), p. 1966-1982.

Sound change in the learner: The perception of connected speech

Andrea Pearman
Universitat Autònoma de Barcelona

Research in speech perception and language learning has posited that acquisition of new, phonetically distinct L2 sounds may be assimilated to existing L1 categories (PAM, Best, McRoberts & Sithole, 1988; Best, 1995) because the L1 phonology acts as a filter, stripping new sounds of L1 phonetically-irrelevant features which may or may not be important phonologically in the L2 (Flege, 1995). As L2 proficiency advances, however, L1 phonetic categories may "evolve" to include phonetic features relevant for both L1 and L2, as postulated in Flege's (1995) Speech Learning Model (SLM). This experiment aims to explore such evolution by studying how learners' phonological expectations change via perceptual readjustment as L2 experience increases.

In connected speech, given particular phonological contexts, listeners are able to backtrack the acoustic consequences of common, productive reduction processes (*e.g.*, assimilation, weakening, deletion, etc.) to arrive at the speaker's intended interpretation. That is to say, upon encountering *izha*, for example, English listeners hold the expectation, borne out of previous experience, that *izha* must be the result of *is + your*. Elaborating on Flege's SLM, it is hypothesized that, when processing reduced speech in the L2, as experience increases, listeners will be more able to interpret common reduction processes in the L2 because of a perceptual readjustment to accommodate L2 relevant (but L1 irrelevant) acoustic information. Thus, more advanced students are expected to interpret cues in the L2 speech signal better than less advanced students.

The present study tests L1 and L2 perception of American English casual speech. A natural English sentence, composed of frequent vocabulary, *Is your friend the one that can't go to bed by ten?*, affected by a mix of reduction processes, *e.g.*, palatalization, place/manner assimilation, vowel/consonant weakening, and deletion, was gated in 80 ms steps. Gates were presented to 12 Americans and 24 Catalans, with two levels of experience: beginner (Cambridge First) and more advanced (Cambridge Proficiency). As in Shockey (1997), after each gate, subjects wrote down what they heard. Subjects' responses were examined in terms of online word recognition: first correct identification for each subject was calculated and "confusions" were analyzed.

In terms of lexical identification, the more advanced Catalan group showed lower, and usually later, recognition peaks than Americans. Less experienced Catalans generally showed no recognition peaks: there was greater dispersion of the results across gates, resulting in no clear peaks. The advanced learners' group recognition totals were consistently lower than natives' and the less advanced learners' group totals were lower than more advanced learners'. Coincident with other studies, learners required more acoustic information than natives in order to identify a word (Nooteboom & Truin, 1980; Koster, 1987), often needing to hear the beginning of the next word before recognizing an item correctly, whereas native speakers tended to recognize the word before its offset.

Regarding the analysis of subjects' confusions, less experienced L2 learners used more signal information, *i.e.*, bottom-up, which led to a greater number of face-value and underanalyzed responses, *e.g.*, the flap in *go to* as [r] and dentalized [nn] in *frien(d) the* and *one that* as [n]'s (in line with L1 phonology). The fact that Catalans relied more on signal

information than on structural or higher level information, led to a lack of grammaticality in final interpretations. Catalans also showed greater difficulty in parsing. Additionally, prosody and predictability played important roles for all three groups.

Results of this study are relevant in underscoring sound change within the learner. As L2 experience increases, learners are seen to perceptually readjust in the direction of L2 phonological processing.

References

Best, C. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), Speech perception and linguistic experience: issues in cross-language research (pp. 171-204). Timonium, MA: York Press.

Best, C., McRoberts, G., & Sithole, N. (1988). The phonological basis of perceptual loss for non-native contrasts: Maintenance of discrimination among Zulu clicks by English-speaking adults and infants. Journal of Experimental Psychology: Human Perception and Performance, 14: 345-60.

Flege, J.E. (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, MA: York Press.

Koster, C. (1987). Word recognition in foreign and native language. Dordrecht, The Netherlands: Foris.

Nooteboom, S., & Truin, P. (1980). Word recognition from fragments of spoken words by native and non-native listeners. IPO Annual Progress Report, 15: 42-47.

Shockey, L. (1997). Perception of reduced forms by native and non-native speakers of English. In J. Leather & A. James (Eds.), New sounds 97. Proceedings of the third symposium on the acquisition of second-language speech (University of Klagenfurt, 8-11 September 1997) (pp. 97-100). Klagenfurt: University of Klagenfurt.

Effects of syllable position on sound change: Aerodynamic and perceptual data on final fricative weakening

Maria-Josep Sole & Andrea Pearman
Universitat Autònoma de Barcelona

Phonetic differences between syllable initial and syllable final consonants have been suggested to be at the origin of some sound changes. In particular, a decreased gesture syllable finally (i.e., less and shorter linguo-palatal contact syllable finally than syllable initially (Byrd 1996; Fougeron, 1999)) has been associated to weakening and/or loss of syllable final consonants (Hock 1986; Recasens 2001). However, the weakening (i.e., gliding, aspiration or loss) of syllable final fricatives in prepausal and preconsonantal position found in a variety of languages may not only be attributed to articulatory reduction, but more crucially to the aerodynamic and perceptual consequences of gestural reduction and of the temporal sequencing of articulatory events. This paper explores the differences in the aerodynamic characteristics for syllable onset and coda fricatives and their perceptual properties. Such differences may account for the weakening of fricatives (i) syllable finally, e.g., gliding (Latin *no**s**, vo**s*** 'we, you' > Italian, Rumanian *no**i**, vo**i***; Latin *men**s**e* 'month' > Occitan [mej]; English *sæ**h**, bisi**g*** > *sa**w**, bus**y***), aspiration (present-day Spanish *de**s**de, do**s*** [h] 'since, two'), and elision (Latin *no**s**, vo**s*** > French *nou**s**, vou**s*** [s] > ø ; English *ye**s*** > *yea**h***; *o**f**, ha**v**e* [v] > ø), and (ii) when followed by a nasal consonant (Latin *me**s**nata* 'kids' > Catalan *ma**i**nada*; Old French *ae(**s**)mer*, Standard Catalan *e**s**ma* > Balearic dialects *e**j**ma,* English *a**i**m;* Old French *ble(**s**)mir*, Provençal *ble**s**mar* > English *blemish;* English *isn't, doesn't, something* pronounced [eɪnt], [ɪnn̩t], [dʌnn̩t] [sʌmn̩] (Gimson 1962)).

It is known that fricatives have tight positional, aerodynamic, and time constraints, vis-à-vis stops, and they allow lesser articulatory and aerodynamic variation than other segment types (Recasens et al. 1997, Solé 2002). If the aerodynamic conditions for generating audible turbulence in coda fricatives are not met, due to a decreased gesture, and/or a reduced oral pressure build up (due to time constraints or a lower rate of flow), fricatives are likely to decay and the resulting sound may be interpreted as a glide or may be perceptually missed. Similarly, in fricative+nasal sequences, the anticipatory velopharyngeal opening for the coarticulated nasal may bleed the oropharyngeal pressure required to produce turbulence for the fricative.

In a first experiment, simultaneous oropharyngeal pressure (Po), airflow and audio-signal were obtained for two American English speakers reading sequences containing symmetrical CVC nonsense syllables, where C=voiced and voiceless fricatives, before a consonant (CVC#C) and phrase-finally (CVC##). The data allow us to observe the timing of aerodynamic and acoustic events, and relate them to inferred articulatory gestures. Aerodynamic and acoustic analyses showed that coda fricatives, vis-à-vis onset fricatives, exhibit (i) a slower oral pressure build-up, (ii) a lower pressure peak, (iii) a delayed onset of audible frication, (iv) a shorter duration (when the syllable is not phrase-final) and (v) a lower intensity of frication (rate of flow being proportional to intensity), compatible with a reduced lingual gesture for coda as opposed to onset fricatives.

The results suggest that a reduced gesture syllable-finally may delay the onset of frication (and thus make it more likely to be affected by overlapping gestures) and may endanger the aerodynamic conditions -- rate and duration of airflow-- for generating audible turbulence,

making the fricative more difficult to detect. A second experiment was carried out in order to test the hypothesis that diminished intensity and reduced duration of turbulence in coda fricatives may lead to the identification of a glide or may be perceptually missed. The CVC stimuli produced by one of the talkers in the first experiment were (i) attenuated by 6 dB step changes, beginning at vowel offset (approximating phrase-final position), and (ii) truncated in 20ms steps beginning at vowel offset (approximating consonant overlap). Subjects were instructed to identify the syllable as CVC, CVj, CVh, CV or to write down whatever they heard, if they heard something else, in normal orthography on the answer sheet.

The results of the perceptual experiment are currently being analyzed and suggest that coda fricatives with diminished intensity of turbulence and/or truncated duration may be reinterpreted as an aspirated [h] -- reflecting the large and slightly turbulent airflow escaping through the open glottis before the fricative constriction has been formed--, an offglide of the vowel, or may be perceptually missed, replicating the historical and synchronic processes.

Finally, the loss and weakening of fricatives followed by a nasal consonant may be attributed to the anticipatory velopharyngeal opening required for the nasal impeding the build-up of the oral pressure necessary for frication. Work by Ohala and Solé (1998) shows that when oropharyngeal pressure during the production of fricatives is vented with catheters simulating the effect of velopharyngeal leakage, fricatives lose much of their high-frequency aperiodic energy, making them frictionless continuants. Our findings suggest that gestural and associated aerodynamic effects may account for the historical and present-day tendency for syllable final fricatives, and for fricatives followed by nasals, to weaken or disappear.

References

Byrd, D. 1996.Influences on articulatory timing in consonant sequences, Journal of Phonetics, vol. 24, pp. 209-244.

Fougeron, C. 1999. Prosodically conditioned articulatory variation: A review. UCLA Working Papers in Phonetics, 97: 1-73.

Gimson, A.C. 1962. An introduction to the pronunciation of English. London: Arnold.

Hock, H.H. 1986. Principles of Historical Linguistics. Mouton: The Hague.

Ohala, J.J. & Solé, M.J. 1998. The controversy of nasalized fricatives. Proceedings of the 135th Meeting of the ICA/ASA. Seattle, Washington, 2921-2922.

Recasens, D., Pallarès, M.D. & Fontdevila, J. 1997. A model of lingual coarticulation based on articulatory constraints," JASA, vol. 102, pp. 544–560.

Recasens, D. 2001. Articulatory and acoustic correlates of syllable position in heterosyllabic consonant clusters. Conference on the Phonetics-Phonology Interface, Zentrum fur Allgemeine Sprachwissenschaft. Berlin (unpublished).

Solé, M.J. 2002.Assimilatory processes and aerodynamic factors," in Laboratory Phonology 7, C. Gussenhoven & N. Warner, Eds., pp. 351–386. Berlin: Mouton.

The effect of pitch accent and word position on the production of vowel sequences: a comparison of Spanish, Romanian, and French

Ioana Chitoran
Dartmouth College

Sequences of high vowel ([i]) and non-high vowel occur in many cognates shared by all Romance languages. In the absence of stress on [i], these [iV] sequences exhibit variation in production, in some languages more than in others. It has been established, for example [1] that in Spanish such vowel sequences tend to be realized as diphthongs (e.g. [bjela] 'rod', [indjana] 'Indiana'), whereas in Romanian the general tendency is to preserve hiatus ([biela], [indiana]). The acoustic parameter compared is the total duration of the vocalic sequence. Shorter duration is taken to indicate a diphthong production [jV], and longer duration, hiatus [i.V].

One of the factors that was found to affect duration in both languages is the position of the vowel sequence in the word. The analysis of data from 4 Spanish speakers and 4 Romanian speakers revealed a word position effect in both languages. [iV] sequences that are initial in the word were found to be significantly longer than word-medial sequences. This finding is consistent with impressionistic accounts by Spanish and Romanian speakers, who report hiatus more frequently when the sequence is word-initial, and diphthongs more frequently when the sequence is word-internal.

The present study focuses on the word-initiality factor, to determine whether it may in turn be affected by the presence or absence of pitch accent in the test sentences. The test words/phrases were elicited in a carrier phrase, read in randomized order, as is customary in laboratory experiments. In both the Spanish and the Romanian sentences the V in the [iV] sequences in test words/phrases bears pitch accent. Examples are given below:

	Spanish	Romanian
word-initial	Dígo con Diana porque sí.	Spune din Diana de trei ori.
word-medial	Dígo mediana porque sí.	Spune mediana de trei ori.
	'I say _____ just so'	'Say _____ three times'

It is therefore unclear whether the vowel sequence duration is sensitive to the boundary of the word as a morphological/syntactic domain, or to the location of the sequence relative to pitch accent. One way of testing the relevance of pitch accent location is by extending the comparison to French, a Romance language which shares the cognates but not the prosodic structure of Spanish and Romanian. Another difference is that French native speakers consistently report the sequences as being produced as diphthongs [jV]. The prosodic structure assumed for French is the one proposed by [2], with a L tone followed by an optional initial rising pitch movement (Hi) in the accentual phrase, and a final H*. The same type of sentences were recorded from three native speakers of French from Lyon:

French word-initial: Dis-nous mes Dianes de nouveau. 'Tell us ____ again'
word-medial: Dis-nous médiane de nouveau.

H* is always realized on the last syllable in the sentence. Hi is optionally realized on the first syllable of the test word/phrase. Therefore, the V of the [iV] sequence does not carry pitch accent. If the word position effect is independent of pitch accent, then it should also be found in French. However, the ANOVA reveals no statistically significant duration difference between word-initial and word-medial vowel sequences for any of the three French speakers. The preliminary conclusion is that the word position effect is crucially enhanced by the presence of

pitch accent (argued also by [3]). These results encourage further comparison of controlled data from different Romance languages, in order to establish a typology of cross-Romance prosodic effects on segmental structure.

References

[1] Chitoran, I. and J.I. Hualde (2002) Variability in hiatus resolution: A phonetic study of [CiV] sequences in two Romance languages. Poster presented at LabPhon 8, Yale University and Haskins Laboratories, June 2002.

[2] Jun, S-A. and C. Fougeron (2002) Realizations of accentual phrase in French intonation. Probus 14, 147-172.

[3] White, L. (2003) English Speech Timing: A Domain and Locus Approach. PhD dissertation, University of Edinburgh.

Intonation as a speech segmentation cue: Effects of speech style

Natasha Warner*, Lynnika Butler* & Takayuki Arai†
*University of Arizona** *and Sophia University*†*

To understand connected speech, listeners must segment it into words. Past research has shown that listeners use many types of information in segmentation (e.g. Cutler et al. 1996, 2002). However, there has been little investigation of whether listeners use intonation in segmentation, and relatively little work on segmentation of differing speech styles. This research addresses these issues, using Japanese. In Tokyo Japanese, there is a rise in pitch at the beginning of each accentual phrase ("APR," accentual phrase rise) (Pierrehumbert and Beckman 1988). Since accentual phrases must begin at a word boundary, listeners might use APRs to help them locate word boundaries. Warner and Arai (2001) found that if listeners hypothesized a word boundary at each APR, this would allow them to locate many of the word boundaries, while producing extremely few false positives (APRs not at word boundaries). A related perceptual study has confirmed that Japanese listeners do make use of APRs in locating words.

Segmentation research has focused largely on careful speech, but speech styles differ widely in ways that may affect the segmentation task. A particular segmentation cue may occur frequently in one speech style but rarely in another. We compared the locations of all word boundaries and all APRs in four Japanese speech corpora: newscasts, a language textbook, spontaneous monologues, and casual phone conversations. We evaluated how often word boundaries are accompanied by an APR. Since a pause is a strong boundary cue, we examined the words which do not follow a pause (the more difficult words) separately. Results appear below:

Speech Style	% non-post-pausal words with APR	% post-pausal words with APR	% all words that follow pauses	% APRs not at word boundaries
Newscasting	44%	98%	36%	0%
Textbook	38%	98%	38%	1%
Monologue	40%	88%	42%	3%
Conversation	37%	68%	32%	1%

Words not after pauses are accompanied by a pitch rise most often in newscasting speech. That is, when there is no obvious boundary cue (a pause), speakers supply pitch rises (an alternate boundary cue) more often in newscasting speech than in other styles. Also, words following pauses almost always have an APR in the more formal speech styles (news and textbook), but not in the less formal styles. Furthermore, monologue speech has the most pauses (because of the demands of speaking fluently with little planning) and conversation the fewest. Finally, very few APRs occur anywhere other than at a word boundary in any style.

These results show that segmentation cues differ in their potential usefulness depending on speech style. Second, speakers provide the intonational word boundary cue most often in exactly that speech style where segmentation is likely the most difficult otherwise (newscasting, with long strings of low-frequency words). Finally, in all speech styles, APRs have the potential to be quite a strong word boundary cue, relative to other known cues (McQueen 1998). Since listeners can listen to, and fluently parse, a great variety of speech styles, it is interesting to determine how speech style affects methods of parsing. Furthermore, investigation of this particular segmentation cue (pitch rises based in the intonational system) brings together the

study of phonetics and phonology of intonation with the psycholinguistic study of speech segmentation.

References

Cutler, Anne, Demuth, Katherine, and McQueen, James. 2002. Universality versus language-specificity in listening to running speech. Psychological science, **13**:258-262.

Cutler, Anne, Norris, Dennis, and McQueen, James. 1996. Lexical access in continuous speech: Language-specific realisations of a universal model. Phonological structure and language processing, ed. by Takashi Otake and Anne Cutler, 227-242. Berlin: Mouton de Gruyter.

McQueen, James. 1998. Segmentation of continuous speech using phonotactics. Journal of memory and language, **39**:21-46.

Pierrehumbert, Janet, and Beckman, Mary. 1988. Japanese tone structure. Cambridge: MIT Press.

Warner, Natasha, and Arai, Takayuki. 2001. Accentual Phrase Rises as a Cue to Word Boundaries. Talk given at the Annual Meeting of the Linguistic Society of America, Washington, D.C., January 2001.

The effect of tone on vowel duration in Thai: A developmental study

Kimiko Tsukada*, Denis Burnham*, Sudaporn Luksaneeyanawin†, Niratasai Krachaikiat† & Sorabud Rungrojsuwan†
University of Western Sydney and Chulalongkorn University-Bangkok†*

It is known that there is a differential effect of tone types on vowel duration such that, other factors being equal, vowels on rising tones are generally longer than those on falling tones and that vowel duration is inversely related to the approximate average fundamental frequency. Although such phonetic phenomena are documented for several languages including Thai and Cantonese (e.g., Gandour, 1977; Kong, 1987) and physiological explanations are offered in the literature, it is not certain if speakers with limited knowledge of the language under investigation (e.g., children or non-native speakers) may exhibit the expected durational patterns. If these phenomena are truly universal and physiologically motivated, they should generalize to all speakers regardless of previous linguistic experience. Alternatively, the interaction between tone and vowel length may be somewhat language-specific and speakers need to acquire this phonetic aspect of the respective language in order to attain native-like proficiency.

In this preliminary study, we attempted to gain an insight into the universality of tonal conditioning of vowel length by examining the speech production of children acquiring Thai as their first language (L1). Three groups of 12 children aged 4, 5 and 6 produced two sets of minimal pairs (/ma:/ and /ja:/) representing four (high, rising, mid, falling) of the five lexical tones of Thai. A maximum of three tokens was elicited from each child in a picture-naming task. Vowel and word duration of each token were measured using EMU speech database system and analysed in separate two-way ANOVAs with Age as a between-subjects factor and Tone as a within-subjects factor.

For both vowel and word duration measurements, only the effect of Tone reached significance [Vowel: $F(3, 93) = 12.7$, $p < 0.001$; Word: $F(3, 93) = 16.0$, $p < 0.001$]. Vowels on a falling tone were significantly shorter than those on the other three tones. This finding is consistent with the results of previous studies (Gandour, 1977; Kong, 1987). However, vowel duration was not differentiated among high, rising and mid tones. Words on a falling tone also had the shortest duration of the four tones. In addition, words on a mid tone were significantly shorter than those on a rising tone. Although the vowels and words became gradually shorter as the children's age increased, the Age effect did not reach significance [Vowel: $F(2, 31) = 0.15$, ns; Word: $F(2, 31) = 0.44$, ns]. This may be due to a substantial inter-subject variability within each age group. The two-way interaction was non-significant [Age x Vowel: $F(6, 93) = 0.86$, ns; Age x Word: $F(6, 93) = 0.52$, ns], indicating that the children of all three age groups had the shortest vowel/word duration for the falling tone.

These results suggest at least two possibilities. First, the association of falling tones with short vowels and, to a less extent, rising tones with long vowels may be acquired relatively early in the children's L1 speech acquisition. Another possibility is that these associations reflect a universal tendency of the production of tone and are expected to be observed regardless of speakers' linguistic experience. We are unable to eliminate the language specificity hypothesis of the interaction between tone and vowel length on the basis of the present results. Our future work includes acoustic analyses of the production data collected from children as young as 18 months old.

References:

Gandour, J. (1977). On the interaction between tone and vowel length: Evidence from Thai dialects. Phonetica, 34, 54-65.

Kong, Q. M. (1987). Influence of tones upon vowel duration in Cantonese. Language and Speech, 30, 387-399.

Prosodic domain effects and vocalic chain shifts

Ewa Jacewicz*, Joseph C. Salmons† & Robert A. Fox*
*Ohio State University** *and University of Wisconsin*†

The relatively recent investigation of the initial edge effects on articulation of consonants initiated by Fougeron & Keating (1997) has provided extensive articulatory and, to some extent, acoustic evidence in support of the view that the effect of "articulatory strengthening" at initial edges of prosodic domains occurs cross-linguistically (Fougeron & Keating, 1997; Keating et al., 1998; Cho & Keating, 2001; Fougeron, 2001; Tabain, 2003a; 2003b; Onaka, 2003). Accordingly, domain edge-initial consonants are produced with increased effort which is manifested in greater articulatory contact. The present work focuses on the first vowel (V_f) immediately following the domain-initial consonant. Systematic investigation of the acoustic characteristics of vowels in this position across prosodic domains has not yet been undertaken.

This study investigates the potential impact of prosodic domain effects in the production and perception of V_f in two Midwestern varieties of American English spoken in Ohio (OH) and Wisconsin (WI), testing the hypothesis that prosodic prominence affects in predictable ways the vowel shifts currently underway in each dialect. In both OH and WI, tense /e/ moves toward the edge of the vowel space and shows some raising, while lax /E/ lowers and/or centralizes. In OH, /e/ shows increased spectral change, while in WI it raises but remains relatively monophthongal. We examined V_f in three distinct prosodic domains: utterance (sentence-initial), phonological phrase (strong branch of a foot), and syllable (weak branch of a foot). In these prosodic contexts, /e, E/ were examined in disyllabic and monosyllabic words in terms of a set of acoustic parameters including vowel duration, formant frequency changes across the vowel, and the extent of formant movement. A perceptual experiment looked at changes in perceived quality of these vowels (removed from their prosodic contexts) in a forced-choice identification task.

The results show that prosodic context has a statistically significant effect on vowel production and perception, and affects vowel acoustics in predictable yet somewhat different ways in each dialect. We also found that prosodic prominence affects "dynamic" vowel properties rather than static ones. Importantly, both speaker and listener are sensitive to this type of variation. For each dialect, vowel duration and the extent of formant movement were influenced by prosodic hierarchy, and the changes across prosodic domains were systematic. The three prosodic contexts yielded opposite patterns of identification: the perceptual accuracy for /e/ increased with higher prosodic prominence whereas the accuracy for /E/ decreased.

The prosodically-conditioned synchronic vowel shifts just described closely parallel one of the best-known patterns in sound change, that of vocalic 'chain shifts'. The core principle of chain shifting, first formulated by (Sievers 1876/81:130-131) but most familiar today as Labov's "Vowel Shift Principle" (1994:262, 601, etc.) predicts: In chain shifts, peripheral [that is, tense] vowels become less open and nonperipheral [i.e., lax] vowels become more open. Vowel systems of American English dialects are at present rapidly diverging from one another, largely along these well-documented paths (Labov 1994, Thomas 2001). These changes in progress provide much finer-grained parallels to our experimental results, as two examples illustrate: First, greater prosodic prominence yields greater dynamic change over time in tense vowels for at least some dialects, while tense vowels in chain shifts are predicted to develop greater offglides (Labov 1994). Second, rather than simply lowering, lax vowels in chain-shifting

dialects appear to centralize, lower and sometimes even front (while lowering). Recent work on chain shifts shows just the same complex pattern for some Midwestern dialects (Gordon 2001).

Our experimental data thus suggest a close connection between these two phenomena and we argue that these ongoing diachronic developments ultimately come from synchronic vowel changes in a given dialect, driven by prosodically-structured variation. Such variation is systematically present for all speakers and listeners in communities undergoing sound change. Prosody is, in short, a key mechanism for segmental change.

References

Cho, T., and Keating, P. A. (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.

Fougeron, C., and Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domain. Journal of the Acoustical Society of America, 106, 3728-3740.

Gordon, M. (2001). Small-town values and big-city vowels: A study of the Northern Cities Shift in Michigan. [= Publications of the American Dialect Society, 84

Keating, P. A., Cho, T., Fougeron, C., and Hsu, C-S. (1998). Domain-initial articulatory strengthening in four languages. In Papers in Laboratory Phonology VI, edited by J. Local, R. Ogden, and R. Temple (Cambridge U. P., Cambridge, England

Labov, W. (1994). Principles of Linguistic Change. 1: Internal Factors. Oxford: Blackwell.

Onaka, A. (2003). Domain-initial strengthening in Japanese: and acoustic and articulatory study. In Proceedings of the 15th International Congress of Phonetic Sciences (Barcelona, Spain), pp. 2091-2094.

Sievers, E. (1881). Grundzüge der Phonetik. Leipzig: Breitkopf & Härtel. (1st ed., 1876.)

Tabain, M. (2003a). Effects of prosodic boundary on /aC/ sequences: acoustic results. Journal of the Acoustical Society of America, 113, 516-531.

Tabain, M. (2003b). Effects of prosodic boundary on /aC/ sequences: articulatory results. Journal of the Acoustical Society of America, 113, 2834-2849

Thomas, E. (2003). Secrets revealed by Southern vowel shifting. American Speech 78.150-170.

The delay of the timing of F0 minima in Akita Japanese

Kenji Yoshida
Shoin University

The present study examines the range of variation of the timing of the F0 minimum in the Akita dialect of Japanese. The aim is to examine whether there is some kind of stability in the alignment of F0 minimum in relation to the phonetic segment, e.g. so called segmental anchoring. The accent system of Akita is classified as a subcategory of Tokyo types, possibly derived from something like Tokyo system with an accentual shift: a phonological change which shifted the position of lexical accent to the following mora. So the other purpose is to explore whether there is phonetic feature which might cause the shift.

The timing of the F0 peak was examined in Yoshida (2003). There are two kinds of F0 peak delay in the dialect, with different magnitude. This delay may be the trigger of the accentual shift. The present study follows up this result and examines, (1) whether there is some delay of F0 minimum, (2) the magnitude of the delay, (3) the influence of the lexical accent position, (4) implication for the tonal specification.

Three native speakers in Nishi-Senboku Town in Akita were chosen from the 31 subjects in Yoshida (2001), who preserved the traditional Akita accent system more than others. The target materials are 3 and 4 syllable nouns (98 and 60 words, respectively). The speakers read the words in two prosodic conditions; utterance final position (F: actually, word spoken in isolation) and non-utterance final position (NF: initially in short sentence). The local minima of F0 where the rise toward the accentual peak starts were measured and normalized with the duration of the accented syllable: the duration of the accented syllable is from 0 to 1.

The results are as following: (1) with no phrasal H as in Tokyo Japanese, the rise starts around the beginning of the accented syllable normally. Although there is pretty much variation due to the micro-prosody, the minima lies from 0 to 0.5 (in normalized measure) in most cases. (2) In NF condition, however, there are the delays of F0 minimum. Moreover, in several cases the magnitude of the delay is about 1; the rise starts at the end of the accented syllable (as in Dominican Spanish: Willis 2003). (3) The larger delays are observed only in the words with the initial or final accented words. In the case of initial accented words, the delay is due to the lack of the time to implement the F0 rise (Xu 2001). In the case of the final accented words, the delay is explained as a suppression of the accent, presumably due to the clash with the H of the following postpositional particle. (4) The result obtained suggests that the beginning of the rise is phonetically specified. F0 fall after the accented syllable is not obligatory (Yoshida 2003b), so the phonological specification of the melody of accent of this dialect is different from that of

Tokyo Japanese, something like L*+H. Therefore, the delay of the minimum rather than maximum should be examined further to explore the cause of the accentual shift.

References

[1] Arvaniti, A., D.R. Ladd, & I. Mannen (1998) Stability of tonal alignment: the case of Greek prenuclear accents, *Journal of phonetics*, 26, 3-25.

[2] ---------- (2000) What is a starred tone? Evidence from Greek., Broe, M. & J. Pierrehumbert (eds.) *Papers in laboratory phonology* 5 Cambridge Univ. Pr.

[3] Ladd, D. Robert & Astrid Schepman (2003) "Sagging transitions" an experimental evidence, *Journal of phonetics* 31, 81-112.

[4] Myers, Scott. (1999) Tone association and f0 timing in Chichewa., *Studies in African linguistics*, 28, 215-239.

[5] Prieto, Pilar, J. van Santen & J. Hirshberg (1995) Tonal alignment patterns in Spanish, *Journal of phonetics*, 23, 429-451.

[6] Silverman, Kim. & J. Pierrehumbert (1990) The timing of prenuclear high accents in English., Kingston & Beckman (eds.) *Papers in laboratory phonology 1*, Cambridge Univ. Press.

[7] Willis, Erik W. (2003) prenuclear low tone alignment in Dominican Spanish., *Proceedings of the 15th ICPhS*, Barcelona. 2941-2944

[8] Xu, Yi (2001) Fundamental frequency peak delay in Mandarin, Phonetica; 58, 26-52.

[9] Yoshida, Kenji (2001) The influence of word familiarity for the standardization of accent. (in Japanese), *Proceedings of the 73rd Meeting of the Dialectological Circle of Japan*, 27-36.

[10] ---------- (2003a) Tonal specification of two types of falling pitch pattern in Akita Japanese., *Proceedings of the 15th ICPhS*, Barcelona, 1293-1296.

[11] ---------- (2003b) Timing control of F0 peak in the dialect of Nishi-Senboku town, Akita. (in Japanese), *Proceedings of the 17th Congress of the Phonetic Society of Japan*, 77-84.

Normalisation of downtrend --- local or global effect?

Ivan Yuen
University of Edinburgh

Downtrend results in different f0 values for a phonologically equivalent accent/tone (Pierrehumbert 1979, Prieto 1996, 1998, Shih 2000). In a late-occurring position-in-utterance, the phonetic f0 value of an accent is realised lower than in an early-occurring position-in-utterance. In equating different f0 values to the same perceived prominence of an accent, it has been shown that listeners compensate for downtrend with regard to a global reference line (Pierrehumbert 1979, Gussenhoven and Rietveld 1988, Terken 1991, Gussenhoven et al 1997). But it is not clear what information in the f0 contour listeners employ in constructing the reference line in order to normalise for downtrend. The previous studies focused mainly on normalisation of downtrend in non-tonal languages, while the current study investigated normalisation of downtrend in a tone language. The use of a tone language allows us to examine which f0 targets in the f0 contour can be employed as a reference frame in normalising for downtrend.

A perception experiment was conducted on Cantonese tones to test whether the listeners compensate for downtrend in tone perception and how different f0 values in different positions-in-utterance can be equated to the same tone. Cantonese has two tones --- High Mid and Mid Low, which are contrastive in terms of f0 height. With downtrend, a late-occurring High Mid tone is realised with a lower f0 level than an early-occurring High Mid tone. The same is also found for a Mid Low tone. On account of downtrend, it is likely that a late-occurring High Mid tone will exhibit an f0 value similar to an early-occurring Mid Low tone. Under such scenario, how can the listener decide whether an f0 value belongs to a High Mid tone or a Mid Low tone?

In the first part of the experiment, subjects were instructed to identify a lexical word occurring in different positions-in-utterance. If the listeners normalise downtrend in equating f0 values to the same tone, it will be predicted that a given f0 value will be perceived as different tones in different positions-in-utterance. Results showed that a given f0 value in a late-occurring position-in-utterance led to more High Level or High Mid tone responses than in an early-occurring position-in-utterance, similar to observations in Pierrehumbert (1979) and Gussenhoven and Rietveld (1988) as found in non-tonal languages.

The second part of the experiment was concerned with whether local f0 context or global f0 context is used as a frame of reference in normalising for downtrend. The local f0 context refers to tone targets immediately adjacent to the test word; whereas the global f0 context refers to all tone targets in the utterance context. Results showed that the listeners did not change their normalisation behaviour, even when the global f0 context was modified, as long as the test word occurred in the same test position-in-utterance and immediately surrounded by the same neighbouring tones. However, when the local f0 context was modified by raising the f0 height of tones immediately adjacent to the test word in a late-occurring position, more Mid Low tones were perceived than when the test word occurred with its original local f0 context.

The results of the experiment will be discussed in terms of time-dependent downtrend and the primacy of local f0 context as a frame of reference in equating different phonetic f0 values to the same phonological tone in Cantonese.

References:

Pierrehumbert, J., (1979), 'The perception of fundamental frequency declination , Journal of Acoustical Society of America, 66(2), 363-369.

Terken, J., (1991), 'Fundamental frequency and perceived prominence of accented syllables', Journal of Acoustical Society of America, 89(4), 1768-1776.

Gussenhoven, C., et al., (1997), 'The perceptual prominence of fundamental frequency peak', Journal of Acoustical Society of America, 102, 3009-3022.

Gussenhoven, C., and Rietveld, A., (1988), 'Fundamental frequency declination in Dutch: testing three hypotheses', Journal of Phonetics, 16, 355-369.

Shih, C., (2000), 'A declination model of Mandarin Chinese' in 'Intonation: analysis, modelling and technology', Kluwer Academic Publisher.

Prieto, P., (1996), ' Pitch downtrend in Spanish', Journal of Phonetics, 24, 445-473.

Prieto, P., (1998), 'The scaling of the L values in Spanish downstepping contours' Journal of Phonetics, 26, 261-282.

Linking of Pitch to Mora: An Experimental Approach.

Shakuntala Mahanta* & K.G. Vijayakrishnan†
Utrecht Institute of Linguistics and Central Institute of English and Foreign Languages-Hyderabad†

It has been reported in the literature (Lahiri and Hayes 1991, Mohanan 1986) that in many languages spoken in India (both of Indo Aryan and Dravidian descent), the pitch melody associated with the non compounded word is an 'lh' contour, with the low and high tones being associated with the left and right peripheries of the word, respectively. In this context, Vijayakrishnan (2001) reports that in Punjabi which has a right edge orientation to word stress (typical of Indo Aryan, Chatterji 1926), in words with no lexical tone and with medial and final stress, the word level pitch remains low till the end of the stressed syllable. In other words, the domain of low tone is from the left edge of the word till the end of the stressed syllable. Taking the assumption that the spreading of low tone to the right is an indication of a rightward shift in stress as our working hypothesis, we examine three related languages which have been claimed to exhibit invariant left edge prominence (Chatterji 1926, Lahiri and Hayes 1991). The languages examined are Odiya, Bangla and Oxomiya (Assamese).

We show with the help of exhaustive experimental evidence conducted on the speech of native speakers of these three languages, that word level prominence is a left-edge oriented phenomenon. Interestingly, statistics show that in Odiya prominence is fixed on the left edge, whereas it shifts to the second syllable in the presence of closed and heavy syllables in Bangla and Oxomiya. Stress shift in Bangla and Oxomiya is restricted to the second syllable only and never goes beyond. To determine this, we took trisyllabic words of the following foot shapes: LLX, HLX HHX, and LHX.

The theoretically viable finding in this experiment is in the nature of stress shift in Oxomiya and Bangla. Throughout our experiment it was noticed that when stress shifts in Bangla and Oxomiya, the low pitch spreads only till the onset of the second syllable. This shows that the expanse of the low pitch is restricted till the head mora and neither the entire foot, nor the whole syllable is linked to it. Statistical significance of the second syllable pitch start of LHX when compared to the pitch start of words with other foot types was in favour of this analysis. Seen from a theoretical perspective, the moraic theory is capable of interpreting this spread as linking of the low pitch to the head mora of the prominent syllable.

Odiya is the strongest contender for fixed left edge prominence, where the low never goes beyond the first mora, and the high tone starts from the second mora itself. Stress shifts in Bangla and Oxomiya to the second syllable, when there is a medial closed syllable. Moreover, this tone needs to be interpreted as a low rise since the low tone restricts itself to the first mora and the upward trend starts immediately.

Our findings may be a pointer in the direction that in these languages the phonetic correlate of prominence i.e., pitch, may be construed at the lowest level of the prosodic hierarchy and not at any level higher than that.

References

Chatterji, S.K. (1926) The Origin and Development of the Bengali Language. George Allen and Unwin Limited.

Hayes, Bruce and Aditi Lahiri (1991) Bengali Intonational Phonology. Natural Language and Linguistic Theory 9: 47-96.

Mohanan, K.P. (1986) The theory of lexical phonology. Dordrecht, Reidel.

Vijayakrishnan, K.G. (2001) Stress and tone in Punjabi. CIEFL Occasional Papers in Linguistics, Vol 10, 2002., Central Institute Of English and Foreign Languages, Hyderabad

The Separation of Pitch and Stress in Onondaga

Heike Lehnert-LeHouillier
University at Buffalo and Haskins Laboratories

This study investigates the separation of pitch and stress in the Northern Iroquoian language Onondaga.

The accentual system of Onondaga has been a long-standing issue in Iroquoian phonology. Stress in Northern Iroquoian languages is said to be phonetically realized by higher F0 and greater amplitude on the stressed syllable (Chafe 1977, Woodbury 1992 for Onondaga). However, unlike the other Northern Iroquoian languages, Onondaga is described to have developed a separation of pitch and stress.

While the phenomenon of separation of pitch and stress in Onondaga has been noted by several researchers (Chafe 1977, Michelson 1988, Woodbury 1992), the exact circumstances under which separation of pitch from stress occurs are somewhat controversial. While Chafe (1977) describes a separation of pitch and stress only for a subset of Onondaga words – primarily those containing the epenthetic vowel /a/, Michelson(1988) and Woodbury (1992) suggest that all words of four or more syllables show this phenomenon.

The current study constitutes the first phonetic investigation of Onondaga accent. Recordings of narratives by two Onondaga speakers are being analyzed acoustically to label F0 peaks and syllables for each word. Each word then is grouped according to F0 peaks with respect to the stressed syllable, numbers of syllables, presence of the epenthetic vowel /a/, and possibly other criteria that may evolve during the analysis. A preliminary analysis of a subset of the data suggests that in addition to some of the previously described conditions for the separation of pitch and stress, pitch placement interacts with vowel length such that pitch is less likely to be high on a long accented vowel and as a result the pitch peak shifts to the pretonic syllable. Furthermore, the preliminary analysis revealed that stress may be marked not only by a high or rising F0 but also by a falling F0. This challanges previous analyses of Northern Iroquoian accentual systems, the Onondaga system in particular as containing two related but independant word prosodic systems: a metrical stress system and a pitch accent system (Hayes 1995, Prince 1983). The fact that the stressed syllable may be marked by a rising as well as a falling F0 in conjunction with the earlier observation that the high pitch only occurs utterance finally (Michelson 1988) suggests that the observed F0 contours may be a sentence prosodic rather than a word prosodic phenomenon.

References

Chafe, W. 1977. Accent and related phenomena in the Five Nations Iroquois languages. In: L.M. Hyman (ed). Studies in Stress and Accent. Southern California Occasional Papers in Linguistics 4.

Hayes, B. 1995. Metrical Stress Theory: Principles and case studies. University of Chicago Press.

Michelson, K. 1988. A comparative study of Lake-Iroquoian Accent. Dordrecht: Kluwer Academic Publishers.

Prince, A. 1983. Relating to the grid. Linguistic Inquiry, 14, 19-100.

Segmental and prosodic factors in a Matbat vowel change

Bert Remijsen
Leiden University

It is well-known that vowel quality intrinsically affects fundamental frequency (F_0), with low vowels having a lower F_0 than high vowels (e.g. Lea 1973, Hombert 1978). As for the reverse relation, that is, the effect of F_0 on vowel quality, various studies have reported that, in singing, F_0 rises are accompanied by rises in the first formant (F_1) (Rossing 1990, Sundberg & Skoog 1997). As far as I know, it remains unclear whether the same phenomenon occurs in normal speech, and, also, whether this intrinsic influence can become extrinsic or phonologized in the language system. In this paper, I present evidence supporting these hypotheses. In the Matbat language, a lowering vowel change is partially conditioned by a rising tone.

Matbat is an Austronesian language with around 600 speakers on Misool, off the west coast of New Guinea. Its immediate ancestor had a five-vowel system with /i,e,a,o,u/ (Blust 1978). I recently found that Matbat has developed a seven-vowel system – /i,e,ɛ,a,o,ɔ,u/ – that is, with an additional level of vowel height. The seven-vowel phonological contrast is evident from lexical minimal pairs and a phonetic investigation based on data from eight speakers. This diachronic change in the vowel system is conditioned by two factors – a phonotactic one and a prosodic one. As for the former, the mid vowels */e,o/ of the original system have lowered when followed by a non-semivowel consonant. For example, /tɛ^{21}l/ 'testicles' has developed from Malay /telur/ 'egg'. We hypothesize that this sound change has started in words that have the coda /l/. Laterals are known to give rise to significant perturbations of the formants of adjacent vowels (Ladefoged & Maddieson 1996). The hypothesis that post-vocalic /l/ started the change is also supported by the fact /l/ conditions lowering among the high vowels as well.

The second factor involved in the vowel change is lexical tone. Matbat has six tonemes (Remijsen 2001). Using a convention in which the speaker's tonal range is represented by the range from 1 (bottom) to 4 (top), these tonemes can be transcribed as follows: /^{41}, ^{3}, ^{12}, ^{121}, ^{21}, ^{1}/. It was mentioned above that the vowel shift only took place in syllables with a non-semivowel coda. A consistent exception to this rule are syllables that have the low-rising toneme (/^{12}/). Words with this toneme that originally had */e,o/ invariably show the shift to /ɛ,ɔ/, respectively, irrespective of the presence or nature of a coda. The rise in F_0 at the end of this toneme implies a corresponding rise in the higher harmonics. I hypothesize that this toneme induced an increase in F_1 as the speakers tuned F_1 upwards so as to follow the nearest rising harmonic. This kind of interaction between F_0 and F_1 ('formant tuning') is well-known from professional singers (Rossing 1990:353, Sundberg & Skoog 1997). A perception study by Carlsson & Sundberg (1992) revealed that formant tuning is not appreciated by expert listeners. This suggests that formant tuning benefits the speaker rather than the hearer, and is therefore likely to be a mechanical, intrinsic characteristic of producing an F_0 rise. From the perspective of the Matbat vowel change process, the rise of F_1 as a result of formant tuning gives open syllables the same kind of transition as induced by a following /l/ in closed syllables. In an inferential test with only two Matbat minimally contrastive word sets, the difference in F_1 at vowel offset between /^{12}/ and the other tonemes tended towards significance (p = .105). We expect that a controlled comparison with a greater dataset will yield significance. Pending such quantitative confirmation, we can already conclude that there is phonological evidence – in the form of this

prosodically conditioned vowel change – that in normal speech, just as in singing, a rise in F_0 may be accompanied by a rise in F_1.

References

Blust, R. (1978) Eastern Malayo-Polynesian: A Subgrouping Argument. In S.A. Wurm & L. Carrington (eds.) Second International Conference on Austronesian Linguistics: Proceedings. Pacific Linguistics Series C61, 181-234.

Carlsson, G. & Sundberg, J. (1992) Formant frequency tuning in singing. Journal of Voice 6, 256-260.

Hombert, J.-M. (1978). Consonant Types, Vowel Quality and Tone. In V.Fromkin (ed.) Tone – A Linguistic Survey. New York: Academic Press.

Lea, W.A. (1973) Segmental and Suprasegmental Influences on Fundamental Frequency Contours. In L.M. Hyman (ed.) Consonant Types and Tones. Southern California Occasional Papers in Linguistics 1, 15-70.

Ladefoged, P. and Maddieson, I. (1996) The Sounds of the World's Languages. Oxford: Blackwell.

Remijsen, B. (2001) Word-prosodic systems of Raja Ampat languages. Utrecht: LOT Ph.D. dissertations series 49.

Rossing, T.D. (1990) The Science of Sound. Reading (Mass.): Addison-Wesley Publishing Company.

Sundberg, J. & Skoog, J. (1997) Dependence of jaw opening on pitch and vowel in singers. Journal of Voice 11, 301-306.

Gradual linguopalatal variations due to a 4-level prosodic hierarchy in French

Yohann Meynadier
University of Edinburgh

A large amount of work has established various kinds of articulatory strengthening of segments and of the modifications of the intra-/inter-segmental articulation dynamics depending on the level of prosodic boundaries. Studies on various languages have showed that a prosodic hierarchy accounts for variations of articulatory gestures and their coordination towards segmental strengthening and of reduction of coarticulation in initial, final or cross-boundary positions of constituents as function as up to 5 hierarchical levels (among others, [1; 2; 6]). In French, studies on this subject mainly focused on accent types [7] or initial position in prosodic constituents of different levels [3].

The present study on French [8] focuses on the effects of a 4-level prosodic hierarchy on intra- and inter-segmental articulatory dynamics of linguopalatal gestures in initial, final and cross-boundary positions of prosodic units. Prosodic-dependant articulatory variations were observed in aC#Ca sequences embedded in a sentence, where CC stands for /kl/, /lk/, /kt/ or /tk/ - # for prosodic boundary -. Accent always fell on the first vowel /a/, which is in the nucleus of the final syllable of the prosodic constituent. 7 to 8 sentences were read without pause 12 or 15 times by 3 French speakers. Prosodic boundaries varied by manipulating the syntactic and thematic structures of utterances. The prosodic hierarchy was composed of 4 levels, from the lower to the higher: syllable/word level < accentual level < non terminal (i.e. continuative) intonation level < terminal (i.e. conclusive) intonation level.

Articulatory analysis of supralaryngeal correlates of the prosodic hierarchy was based on an EPG investigation of the temporal and spatial dimensions of linguopalatal gestures in aC#Ca. 178 different EPG measurements were extracted for each sequence, related to the articulation of the vowels and the consonants, as well as to the VC, CC and CV coarticulation inside the sequences. To extract the linguopalatal correlates of the prosodic hierarchy, a multi-step selection procedure of articulatory results was done. It was mainly based on two statistical criteria: a minimal narrow correlation with the prosodic hierarchy design (calculated by means of the Spearman rank correlation coefficient), and a minimal homogeneity between speaker behaviours (i.e. at least 2 out the 3 speakers must show the same direction of articulatory variation for the same EPG measurement).

The results showed that 10-14 % of the different EPG measurements constituted the core of articulatory co-variations which increased inter-individually and gradually according to the prosodic hierarchy. In this study, the higher the prosodic level, (1) the more the pre-boundary open vowel a_1 showed a large duration and amplitude of its opening; (2) the more the final rime aC# showed a narrow inter-gestural cohesion; (3) the more C#C coproduction was reduced; (4) the more the cross-boundary articulatory timing between C_1 and a_2 was large.

Although these results did not reproduce Fougeron's on French, regarding the gradual articulatory strengthening of maximal linguopalatal constriction of segments in initial prosodic positions, they closely showed the same frequencies of hierarchical distinctions. In both studies, the prosodic-dependant articulatory correlates essentially accounted for 3 different strengths of hierarchical distinction between prosodic levels, which are in decreasing order: intonational level (terminal or non terminal) *versus* non intonational level (accentual and syllable/word levels) >

accentual level *versus* syllable/word level > terminal intonational level *versus* non terminal intonational level.

The interpretation of the results concerns the architectures of the phonological representation based on a hierarchical prosodic constituency of speech structure. The discussion focuses mainly on the nature and, the number of, as well as the relationships between prosodic hierarchical levels, determined by specific intonational models of French prosody [4; 5].

References:

[1] Byrd, D., Kaun, A., Narayanan, S. & E. Saltzman (2000). Phrasal signature in articulation. In M.B. Broe & J. B. Pierrehumbert (eds), Papers in Laboratory Phonology V, 70-87. Cambridge: CUP.

[2] Cho, T. (2001). Effects of Prosody on Articulation in English. UCLA: PhD dissertation.

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

[4] Hirst, D. & A. Di Cristo (1984). French intonation: a parametric approach. Die Neueren Sprache, 5, 554-569.

[5] Jun, A-S. & C. Fougeron (2000). A phonological model of French intonation. In Intonation: Models and Technology, A. Botinis (ed.). Dordrecht: Kluwer, 209-242.

[6] Keating, P., Cho, T., Fougeron, C. & C.H. Hsu (2003). Domain-initial articulatory strengthening in four languages. In J. Local, R. Odgen & R. Temple (eds), Papers in Laboratory Phonology VI. Cambridge: CUP.

[7] Loevenbruck, H. (2000). An investigation of articulatory correlates of the accentual phrase in French. Proceedings of the 14th ICPhS, 667-670. San Francisco.

[8] Meynadier, Y. (2003). Interaction entre prosodie et (co)articulation linguopalatale en français. Université de Provence: PhD dissertation.

Towards an Exemplar-Based Phonological Model of the Perception of Dialect Variation

Cynthia G. Clopper
Indiana University

Phonological change does not occur in a vacuum, but instead is crucially associated with the identification of a speech community (Labov, 2001). Similarly, synchronic phonological variation is crucially linked to the social identity of the talkers. Sociolinguists explore language variation in terms of both phonological variants and the social categories associated with different forms (Labov, Ash, & Boberg, in press; Thomas, 2001). The perception of phonological categories has also been studied extensively since the categorical perception of phonemes was first demonstrated (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). However, the perception of social categories has only recently been empirically investigated. In a series of studies, Clopper and her colleagues (Clopper, Conrey, & Pisoni, 2003; Clopper & Pisoni, 2003, in press a, in press b) examined the perceptual categorization of dialect variation. Their findings indicate that just as naïve listeners have stable and robust mental representations of phonological categories, they also have stable and robust mental representations of dialect categories.

One approach to modeling perceptual categorization is the exemplar view, in which all events are encoded and stored in memory. Bybee (2001) and Pierrehumbert (2001, 2002) have proposed exemplar models to account for lexical frequency effects in phonological change. One crucial aspect of their models is a strong link between phonological and lexical representations. In the same way, an exemplar model that includes links between phonological and talker-specific representations may be able to account for the perception of phonological variation. Just as an exemplar model provides a mechanism for comparing novel utterances with previously encountered utterances to determine the semantic content of the message, an exemplar model could also provide a mechanism for comparing a novel utterance produced by an unfamilar talker with previously encountered utterances to determine the dialect of the talker.

Evidence that naïve listeners perceive the structure of phonological systems was provided by Singh and Woods (1970), who asked naïve listeners to judge the similarity of pairs of American English vowels. A multidimensional scaling analysis of the listeners' judgments revealed a structured perceptual similarity space with two prominent dimensions that correspond to the articulatory dimensions of tongue height and advancement. Similarly, Clopper and her colleagues (Clopper et al., 2003; Clopper & Pisoni, 2003, in press b) have repeatedly uncovered a perceptual similarity space of American English regional dialects with two prominent dimensions that reflect the phonological differences between the vowel systems of the different regional varieties. In addition, Clopper and Pisoni (in press a) reported that participants who had lived in a given region more accurately categorized talkers from that region than participants who had lived in other regions. Their findings suggest that experience is dialect-specific, providing further evidence for an exemplar model where representations are composed of actual experiences of events. Experience with variation does not lead to an overall greater sensitivity to variation but instead reflects detailed knowledge about the specific varieties that an individual has previously encountered.

An exemplar-based model of phonological variation can thus account for the perception of social structure as well as linguistic structure. Just as individual phonemes and lexical items must be identified when listening to speech, the social identity of the talker can also be identified through the comparison of the novel utterance to existing exemplars of linguistic varieties stored

in memory. An exemplar model provides a single mechanism for processing both linguistic and non-linguistic information in the acoustic speech signal and can account for the perceptual similarity structure of phonological systems and social identity.

References

Bybee, J. (2001). Phonology and Language Use. Cambridge: Cambridge University Press.

Clopper, C. G., Conrey, B. L., & Pisoni, D. B. (2003). Effects of gender on the perceptual categorization of regional varieties of American English. Paper presented at the 48th Annual Meeting of the International Linguistics Association, New York, NY.

Clopper, C. G., & Pisoni, D. B. (2003). Free classification of regional varieties of American English. Poster presented at New Ways of Analyzing Variation 32, Philadelphia, PA.

Clopper, C. G., & Pisoni, D. B. (in press a). Homebodies and Army Brats: Some effects of early linguistic experience and residential history on dialect categorization. Language Variation and Change.

Clopper, C. G., & Pisoni, D. B. (in press b). Some acoustic cues for the perceptual categorization of American English regional dialects. Journal of Phonetics.

Labov, W. (2001). Principles of Linguistic Change: Social Factors. Malden, MA: Blackwell.

Labov, W., Ash, S., & Boberg, C. (in press). Atlas of North American English. Berlin: Mouton de Gruyter.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. Psychological Review, 74, 431-461.

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition, and contrast. In J. Bybee and P. Hopper (Eds.), Frequency and the Emergence of Linguistic Structure (pp. 137-157). Amsterdam: John Benjamins.

Pierrehumbert, J. B. (2002). Word-specific phonetics. In C. Gussenhoven and N. Warner (Eds.), Laboratory Phonology 7 (pp. 101-140). Berlin: Mouton de Gruyter.

Singh, S., & Woods, D. R. (1970). Perceptual structure of 12 American English vowels. Journal of the Acoustical Society of America, 49, 1861-1866.

Thomas, E. R. (2002). Sociophonetic applications of speech perception experiments. American Speech, 77, 115-147.

Online Use of Phonetic Detail in Spoken Word Recognition

Katherine Crosswhite & Michael K. Tanenhaus
University of Rochester

We report on two eyetracking studies demonstrating that listeners are sensitive to fine-grained phonetic detail during online spoken word recognition. The first shows that listeners differentiate onset-embedded word pairs (e.g. *ham-hamster*) before reaching phonemic disambiguation. Previous research suggests use of vowel duration to predict word length (Salverda et al. 2003, Crosswhite et al. 2003., Davis 2000). However, experiment 1 shows that disambiguation occurs very early, before the end of the 1st vowel. Experiment 2 investigates use of formant trajectory slope, which is correlated with duration but assessable early in the vowel. Results show recognition is delayed when formants are made initially similar to the competitor word (i.e., made steeper or shallower) but where vowel duration is unchanged.

In experiment 1, participants heard prerecorded commands like "Click on the hamster" or "Click on the ham". 16 easily-picturable embedded-word pairs were used; each participant heard only one of the possible sentences for each pair. Trials commenced with 500 ms of exposure to 4 pictures on a computer monitor, followed by presentation of the corresponding stimulus sentence. Trials ended when the subject clicked an object. In experimental trials, one picture corresponded to the target word, one to the competitor, and the other two were phonetically unrelated to either target or competitor. In 16 filler trials, an embedded-word picture pair was present but not referred to, and in 16 all four were unrelated. Participants' eye gaze was monitored using an SMI Eyelink head-mounted eyetracker with a sampling rate of 250 Hz. Point of disambiguation between target and competitor was evaluated by calculating the proportion of fixations to each object in every 4 ms time slice. Results show that fixations to both target and competitor far exceed that for the unrelated objects once the initial portion of the word is heard. Fixations to the target instead of the competitor predominate by the end of the word as participants click on the correct object. This rise begins during the first vowel of the word, with target and competitor already differentiated 100 ms before the end of the vowel.

Experiment 2 used the same visual display as experiment 1, but experimental stimuli were subject to PSOLA resynthesis in Praat. For monosyllables, the 1st half of the vowel was decreased in duration to 70%, with the 2nd half increased to 130%, maintaining the overall duration; this produced steeper C-to-V formant transitions. Since formant steepness is inversely correlated with duration (Moon and Lindblom 1994), formants in the initial portion of the vowel were more consistent with those typical for the competitor disyllable. A similar process was applied to the disyllables, except the 1st half was lengthened and the 2nd half shortened, producing shallower transitions. Additional filler items with non-resynthesized embedded word pairs were also included to preclude training effects. Results show that changing formant steepness hindered recognition. For monosyllables, although differentiation of target from competitor still occurred before the end of the vowel, it was delayed by about 50 ms. For disyllables, even later disambiguation, nearly 100 ms after the end of the vowel, occurred. This is consistent with the fact that formant trajectory cues in the 2nd portion of the vowel remain consistent, perhaps even exaggeratedly so, with the target: because the second portion is shortened in disyllables but lengthened in monosyllables, resynthesized disyllables should be more ambiguous than the resynthesized monosyllables, in keeping with the results obtained.

These results add to the growing evidence suggesting that listeners attend to seemingly unimportant fine phonetic detail, and that this can impact offline judgments (cf. Goldinger 1996, West 1997). Results of the current studies suggest that formant steepness has a similar status, and that phonetic detail is used in online speech processing.

References

Crosswhite, Katherine, Mikhail Masharov, Joyce McDonough, and Michael K. Tanenhaus. 2003. Phonetic Cues to Word Length in the Online Processing of Onset-Embedded Word Pairs, Ms., University of Rochester.

Davis, Matt. 2000. Lexical Segmentation in Spoken Word Recognition. Ph.D. Thesis, Birbeck College.

Goldinger, S.D. 1996. Words and voices: Episodic traces in spoken word identification and recognition memory. Journal of Experimental Psychology: Learning, Memory and Cognition 22, 1166-1183.

Salverda, A.P., D. Dahan & J. M. McQueen. 2003. The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. Cognition 90, 51-89

West, P. 1997. Perception of distributed co-articulatory properties of English /l/ and /r/. Journal of Phonetics 27, 405-426.

Isolating the cognitive processes that underlie disfluencies and phonological form

Michelle L. Gregory*, Julie Sedivy‡, Anjula Joshi† & Danial Grodner‡
*University at Buffalo**, *Brown University*‡ *and Univeristy of British Columbia*†

Disfluent speech accounts for approximately 6% of spoken language. In production, disfluencies have been hypothesized to be the result of planning difficulties. However, there have been relatively few studies that have been able to isolate the cognitive processes that result in disfluent productions. In this study we tested the hypothesis that disfluencies and word duration are the result of increased processing difficulty by investigating the disfluencies preceding prenominal adjectives with varying processing loads. Given that word duration and disfluencies are correlated with planning problems, it was hypothesized that speakers would be more likely to produce longer adjectives and more disfluencies (a) in the presence of a contrast, and (b) in the presence of adjectives that require a comparative judgment (scalar adjectives, such as "small" "tall", etc.) over adjectives that denote an inherent property of the object itself, such as a color or material.

To test the interaction between processing load, length, and disfluencies, we analyzed spontaneously generated speech from sixteen speakers in an elicited production task in which a speaker instructed a hearer to move one of the objects in the display. On half of the trials, the target object was the only object of its kind, whereas on the other half, a contrasting object of the same category was present. The distinguishing property for the contrast trials was either color, material, or a scalar property such as height, width, etc. The adjectives were coded for duration and adjacent disfluencies (repetitions, restarts, filled and silent pauses). To determine whether disfluencies are correlated with the increased processing demand of denoting a contrast set, comparisons were made between the color adjectives uttered in a contrast set and those uttered in a non-contrastive context. Color adjectives that denote a contrast do not have a greater rate of disfluencies than those that denote non-contrast; $p>.1$. Additionally, despite having an increase in pitch range, color adjectives which denoted a contrast were not longer than those in a non-contrastive environment, $p>1$. To determine whether comparative judgments increased length or disfluency rates, comparisons were made between scalar adjectives to material and color adjectives in a contrast set. There were more disfluencies with the use of scalar adjectives than either material; $p<.01$, or color adjectives; $p<.01$. Similar results were found for duration; scalar adjectives were longer than color and material adjectives, $p< .02$. These results demonstrate that denoting a contrast set does not increase processing difficulty, but comparative judgments do increase processing load during production, as measured by disfluency rates and length.

We used eye tracking data to isolate the origin of the increase in length and disfluency rate. Eye movements were recorded for 5 regions; (e.g., uh, can you move uh the red cup)--an initial disfluency (uh), preamble (can you), the verb (move), determiner (the), and the adjective (red). An anova by subject revealed that there is an increase in the time looking at the contrasting object in both the determiner region, $F(2,20)= 4.960$, $p<.05$, and the adjective region, $F(2,20)= 3.395$, $p<.05$ for scalar adjectives when compared to color or material adjectives. There was no difference in looks to the target object. These data demonstrate that the increase in duration and disfluency rate with scalar adjectives is solely due to an increase in time spent encoding the adjective and not the target object itself. In sum, speakers take to longer produce scalar adjectives, and produce them with more disfluencies, than other adjective types. Eye tracking data confirm that speakers spend more time looking to the contrasting object directory preceding

and during the production of the adjective. These facts suggest that the increased processing difficulty arises from the speaker encoding the appropriate relative property to distinguish the two objects.

Recursive patterns of phonological change in the lexicon

Michele L. Morrisette
Indiana University

When children begin acquiring a new sound, they often produce the sound correctly in some words, but not others. The resultant gradual process of sound change, known as lexical diffusion, raises an interesting question of why some words appear to be more susceptible to change than others. Recent investigations in acquisition have linked sound change to the lexical characteristics of words, e.g., word frequency, neighborhood density, age of acquisition, grammatical category (Gierut & Storkel, 2002; Morrisette, 1999; Tyler & Edwards, 1993). This paper reports findings from 2 studies that reveal that sound change follows a cyclic pattern, being implemented recursively in words of high then low frequency.

The purpose of the first investigation was to cross-sectionally examine the course of sound change for 20 preschool children in the process of acquiring coronal and velar stop consonants. Importantly, children were at different stages in the process of acquiring the stops. That is, some children produced the stops accurately in only a few words that were sampled, whereas other children produced the stops accurately in many words that were sampled. Rank ordering the children on the basis of the number of words with accurate stop productions provided a continuum of learning for these sounds. For each child along the continuum, the average frequency of words produced accurately with the target stops was then calculated and compared to the average frequency of words produced inaccurately. Results revealed that as children progressed from accurate productions in few to many words, they alternated between changing high and low frequency words in the lexicon.

The purpose of the second investigation was to validate this recursive pattern of sound change in words longitudinally for 2 preschool children acquiring velar stop consonants. These 2 children's productions of velar stops were sampled approximately every 2 weeks over the course of 3 months. The average frequency of words produced accurately with the target velars was calculated and compared to the average frequency of words produced inaccurately at each point in time. Results from this longitudinal investigation supported findings from the cross-sectional investigation. Both children acquired the velar stops by alternating between high and low frequency words across time. What differed across children was the start of the recursive cycle. One child evidenced change first in relatively high frequency words, while the other children evidenced change first in relatively low frequency words. Nonetheless, the recursive alternation between change in high and then low frequency words was uniformly predictable.

While the role of cyclicity is novel to the study of sound change in words, similar cyclicity effects have been documented for featural development in both normal and disordered phonological acquisition (e.g., Gierut, 1996; Gierut & O'Connor, 2002; Ingram, 1990), as well as in other areas of development (Smith & Thelen, 1993). A parallel distributed processing framework (Smolensky, 1987) holds promise in accounting for these results. In particular, it is argued here that children strive to achieve harmony in the lexicon by selecting words of alternating frequencies to undergo sound change. As high frequency words in the lexicon are activated in sound change, the child then strives for lexical harmony by activating low frequency words. Once low frequency words are activated for change, the child then seeks the reverse activation of high frequency words. The result is a recursive pattern of sound change in words of the lexicon.

References

Gierut, J. A. (1996). An experimental test of phonemic cyclicity. Journal of Child Language, 23, 81-102.

Gierut, J. A., & O'Connor, K. M. (2002). Precursors to onset clusters in acquisition. Journal of Child Language, 29, 495-517.

Gierut, J. A., & Storkel, H. L. (2002). Markedness and the grammar in lexical diffusion of fricatives. Clinical Linguistics & Phonetics, 16, 115-134.

Ingram, D. (1990, November). The acquisition of the feature [voice] in normal and phonologically delayed English children. Paper presented at the American Speech-Language-Hearing Association Convention, Seattle, WA.

Morrisette, M. L. (1999). Lexical characteristics of sound change. Clinical Linguistics & Phonetics, 13, 219-238.

Smith, L. B., & Thelen, E. (1993). A dynamic systems approach to development: Applications. Cambridge, MA: MIT Press.

Smolensky, P. (1987). Information processing in dynamical systems: Foundations of harmony theory. In D. E. Rumelhart, J. L. McClelland and the PDP Research Group (Eds.), Parallel distributed processing: Explorations in the microstructure of cognition (Vol. 1, pp. 194-281). Cambridge: MIT Press.

Tyler, A. A., & Edwards, M. L. (1993). Lexical acquisition and acquisition of initial voiceless stops. Journal of Child Language, 20, 253-273.

Distinguishing Speaker and Linguistic Variation Using Configurable Articulatory Synthesis

Khalil Iskarous, Heike Lehnert, Louis Goldstein & Mark Tiede
Haskins Laboratories

The assumption of universal phonetic categories has been called into question in a number of recent works (Pierrehumbert 2000, Fleming 1997, Pierrehumbert et al. 2001). A variety of experiments have shown that the phonetic details of the units of contrast can differ from one language to another (Keating 1984, Bradlow 1995). Bradlow 1995, for instance, showed that speakers of English and Spanish vary systematically in the phonetic details of "the same" vowel. But since the variation in the data could be due to speaker differences, Bradlow explicitly eliminates this possibility. These two sources of variation are always met in experiments on phonetic detail: biological variation between speakers and cultural variation in how the articulators are used by speakers of the same language community. In most work in phonetics, speaker variation is treated like token-to-token variation and experiment design techniques are used to eliminate it from consideration. But biological variation between speakers is certainly present along with cultural variation in the speech that the child learns from. There is indeed evidence from speech perception research that speaker differences are not only remembered, but play an important role in word perception (Pisoni 1998). It is therefore important to model both biological and cultural variation, and to see how they are integrated and distinguished. Two approaches are possible: an acoustic signal approach or a speech production systems approach. In the signal approach, the two types of variation would be studied with respect to their affect on the formant frequencies, VOT, segmental durations etc., whereas in the systems approach, the variation would be studied with respect to biological variables like size of vocal tract, lip shape, etc. and phonetic variables like control of tongue shape and lip motion strategies. We take a systems approach here, since little is known about how acoustic parameters actually differentiate between speakers. Form this perspective, phonetic detail of phonological contrasts is seen as a specific way in which speakers of the same language community control articulators that systematically vary from one speaker to another. In this work, we model the two types of variation using CASY (Configurable Articulatory Synthesis), a speech synthesis system that generates sound from a time varying model of vocal tract geometry (Rubin et al. 1981, Rubin et al. 1996). Two sets of parameters are controllable: one set refers to the positions and shapes of various articulators and the other set determines the shape and size of the hard structures like the hard palate and posterior pharyngeal wall. These parameters can be set by fitting the model articulators to MRI, EMA, and ultrasound data. CASY can therefore be used to systematically alter the size and overall shape of the vocal tract to simulate different speakers and at the same time to simulate different strategies of phonetic detail. We will report data from an experiment on American English /r/. The reason for this choice is that speaker variation involved in this segment isn't only due to different sizes of vocal tracts, etc. but is also due to different articulatory strategies used by different speakers. Ultrasound and acoustic data of /r/ produced in three contexts was collected from 6 native speakers of American English and the articulatory data was then used to generate CASY parameters. Speaker variation could be detected in whether bunching or retroflexing was used. But a certain aspect of phonetic detail was consistently controlled by the speakers: whichever part of the tongue was used by different speakers, constrictions in the palatal, labial, and pharyngeal regions were achieved. This is

consistent with the findings of Delattre and Freeman (1968) and Gick and Goldstein (2002). CASY manipulation of the constrictions at the primary gestures for differently shaped vocal tracts and perceptual judgments of the simulations further distinguishes between aspects of the data due to speaker variation and those due to the cultural use of the tongue to distinguish American English /r/ from other varieties.

Changes in articulatory settings and aerodynamic conditions as sources of sound change

Didier Demolin
Universidade de Sao Paolo and Université Libre de Bruxelles

Changes in the temporal coordination of articulators, necessary to produce segments, can trigger modifications of the aerodynamic conditions in the vocal tract that can themselves provoke different or new acoustic outputs. These new acoustic outputs, if not normalized by listerners, can then be the source of sound change. The paper shows, based on aerodynamic, articulatory and acoustic data, how such changes are possible from examples taken in various languages. Ohala (1983) and Yigezu (2001) showed how the interaction between articulatory settings and aerodynamic conditions can explain the origin of implosives from geminate consonants. Denning (1989) showed the same things to explain the diachronic development of voice quality. The data studied in this paper give examples observed in various languages of the world. These changes involve the temporal coordination and control of articulators as well as overshoots in the targets of articulators. One case of a change in aerodynamic conditions and its effect on articulation will also be examined.

The origin of the complex labio-uvulars consonants in Mangbutu-Efe languages can be explained by changes in the articulatory setting and aerodynamic conditions of the vocal tract when speakers produce consonants involving double closures. Indeed the sound change /gb/ > /qb/, in which the second complex consonant combines voiced and voiceless features, can be explained by a combination of the loss of the posterior contact in the oral cavity and changes in the aerodynamic conditions in the vocal tract, both in the oral and in the pharyngeal cavity. In a first phase the lowering and backward movement of the tongue, made during the double closure, triggers a lowering of the pressure in the oral cavity. At the same time, pressure increases in the pharyngeal cavity with adducted vocal folds. Then the rapid balance between pharyngeal and sub-glottal pressure inhibits the voicing that is clearly distinguishable at the beginning of the consonant. When the posterior closure is released because of the backing movement of the tongue, the airflow going from the high pressure in the pharyngeal cavity to the low pressure oral cavity sets the vocal folds in vibration again. There is therefore no opening and closure of the vocal folds in the double stop. The temporal coordination of closure releases is also crucial to explain an allophone of the voiceless labio-vlear stop /kp/, i.e. [kB] in which the release sounds as a trill. However there is no real trill in this case, there is just a different coordination of the velar and closure release, which is first velar and then labial in the case of /kp/ and labial and then velar in the case of [kB] The auditory impression of a trill is given by the 2 consecutive bursts. The origin of the palatal unexploded implosive consonants found in Hendo can be explained by similar factors (Demolin et al. 2001). Similar changes have been observed synchronically in Karitiana (Storto 1999). Tupi languages also provide good examples of articulatory and aerodynamic interactions triggering changes with the final unreleased stops found in these languages.

The origin of post-alveolar or retroflex stops in some languages of Africa (e.g. Surmic and Fon) can be explained by the maintaining of the place of articulation of sounds often described as alveolar implosives. In this case, the loss of the initial articulatory and aerodynamic conditions, i.e. the loss of vocal folds adduction and of an ingressive airflow produce a post-alveolar or retroflex consonant which corresponds to the place of articulation of the original implosive.

References

Demolin, D., H. Ngonga et A. Soquet (2002). Phonetic Characteristics of an Unexploded Palatal Implosive in Hendo. Journal of the International Phonetic Association, 32, 1. 1-17.

Denning, K. (1989). The diachronic development of voice quality. Ph. D. Disseration. Stanford University.

Ohala J.J. (1983). The Origin of Sound Patterns in Vocal tract Constraints. In The Production of Speech. (P. F. MacNeilage editor), pp. 189-216. New-York. Springer-Verlag.

Storto, L. (1999). Aspects of Karitiana grammar. PhD Dissertation MIT.

Yigezu, M. (2001). A comparative study of the phonetics and phonology of Surmic languages. PhD Dissertation, Université Libre de Bruxelles.

Discourse-level effects on durations in synthesized English and French: testing listener preferences

Caroline L. Smith
University of New Mexico

Many researchers have found evidence that speakers modify several dimensions of their speech as a function of the discourse structure of the material being spoken (e.g., Fon, 2002; Grosz and Hirschberg, 1992; Hirschberg and Pierrehumbert, 1986; Lehiste, 1975; Swerts and Geluykens, 1994; van Donzel, 1999). Dimensions in which such an effect has been demonstrated include pitch contour, pitch range, RMS amplitude, speech rate, final lengthening and pause duration. From a linguistic perspective, evidence that discourse-level organization has a measurable effect on production implies that prosody must reflect every level of linguistic structure. However, these patterns are also important for application to speech synthesis and recognition. For synthesis, naturalness and intelligibility can benefit from discourse-appropriate prosodic modifications. This requires that discourse-level patterns pleasing to listeners must be identified , and they must be at least potentially implementable from an automatic analysis of discourse organization. The study presented here tests some durational modifications that English listeners liked in natural speech (Smith in press), and investigates whether they are also perceived as improvements in synthesized speech. This study also compares the perceptions of English and French listeners; previous results suggest that speakers of these languages differ in the magnitude of the discourse-level effects on timing that are produced (Smith, Hogan and McCraw 2003).

For each language, an instructional text that had been previously analyzed for a production study was used as the basis for this study. In each text, the transition between each sentence and the next was classified as belonging to one of four topic transition types (Shift, Continuation, Elaboration or Marker of text structure). The classification was the majority opinion of four (French) or five (English) linguistically-trained native speakers. From each of the texts, three shorter passages were extracted for synthesis in Festival for English and MBROLA for French.

Three different synthesized versions were created for each of the three shorter passages. For English, the base version takes its durations from the defaults provided by Festival. For French, the implementation of MBROLA being used requires that durations be explicitly supplied for each segment; base version durations are calculated from the global means of the durational parameters measured in the production study. The other, modified versions are calculated identically for the two languages. The mean version modifies the base version differently for each topic transition type by setting pause duration, amount of final lengthening and speech rate equal to their mean values calculated across all sentences in the entire text that were classified with that type of transition. The pause version uses pause durations sampled from the values measured in the production study, using values produced with the appropriate topic transition type. Final lengthening and speech rate are the same as in the mean version.

Listener evaluation will follow a protocol similar to that of Smith (in press). The various modified versions of the English passages will be presented to ten listeners who are native speakers of English, and the French passages to ten French listeners, in a forced-choice pairwise comparison task. Listeners will be seated in front of a computer, and will be able to control their rate of progress through a series of screens making up the experiment. Each screen will present

two different versions of a passage, and listeners will click to select which version they prefer. On the basis of the results obtained with modified versions of natural English speech, it is expected that the mean and pause versions will receive higher ratings than the base version, which would indicate a preference for versions which include discourse-level patterns.

References

Fon, Y-J. J., 2002. A Cross-Linguistic Study in Syntactic and Discourse Boundary Cues in Spontaneous Speech. Ph.D. dissertation, The Ohio State University.

Grosz, B., Hirschberg, J., 1992. Some intonational characteristics of discourse structure. In: Proceedings of the 2nd International Conference on Spoken Language Processing, Banff, pp.429-432.

Hirschberg, J., Pierrehumbert, J., 1986. The intonational structuring of discourse. In Proceedings of the 24th Annual Meeting of the Association of Computational Linguistics, pp. 136-144.

Lehiste, I., 1975. The phonetic structure of paragraphs. In Cohen, A., Nooteboom, S. G. (Eds.), Structure and Process in Speech Perception. Proc. ? Symposium on Dynamic Aspects of Speech Perception. Springer-Verlag, New York, pp. 195-203.

Smith, C., in press. Topic transitions and durational prosody in reading aloud: production and modeling. Speech Communication.

Smith, C., Hogan, L., McCraw, M., 2003. Reading aloud a connected text: how the organization of topics affects sentence-final lengthening in English, French and Japanese. Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona, Spain, pages 1799-1802.

Swerts, M., Geluykens, R., 1994. Prosody as a marker of information flow in spoken discourse. Language and Speech 37, 21-43.

van Donzel, M., 1999. Prosodic Aspects of Information Structure in Discourse. Thesus, The Hague.

Festival: http://www.cstr.ed.ac.uk/projects/festival/
MBROLA: http://www.tcts.fpms.ac.be/synthesis/

Experimental evidence and the nature of the schwa/zero alternation in French

Marie-Hélène Côté* & Geoffrey Morrison†
Université d'Ottawa and University of Alberta*†

Schwa/Ø alternations in French appear in two broad contexts:

•morpheme-internally (*cerise* "cherry" [s(œ)riz])

•at various types of junctures, in particular word boundaries (*valse rapide* "fast waltz" [vals(œ)rapid] and clitic boundaries (*ce bateau* "this boat" [s(œ)bato]).

These alternations have been variously analyzed in terms of schwa epenthesis or deletion. It is accepted that morpheme-internal schwas must be underlying given their contrastive nature in the lexicon (compare *pelouse* "lawn" [p(œ)luz] with *blouse* "blouse" [bluz] *[bœluz]). But junctural schwas are never contrastive (Tranel 1981; Côté 2000) and their presence is determined by segmental and prosodic factors. This suggests that they should be considered epenthetic and not lexically present. Yet this claim is challenged by experimental evidence suggesting the underlying nature of schwas at clitic boundaries. It is claimed here that such evidence is disconfirmed by additional experimental results or should be reanalyzed without reference to underlying schwas. We maintain that *all* junctural schwas are epenthetic.

Phonetic studies have identified two facts which are interpreted as supporting the underlying nature of schwas in clitics. They are considered in turn, and shown to be inconclusive.

I. Some lip rounding can be observed in the vicinity of the clitic [k] in (1), while no such rounding appears in an identical context not involving schwa (2) (Barnes & kavitskaya 2002). Lip rounding is interpreted as originating from an underlying schwa (which corresponds to a front rounded vowel [œ] or [ø]).

(1) *plus p'tit qu'la femelle* "smaller than the female" [ply pti k#la fœmel]

(2) *tout p'tit clavecin* "really small harpsichord" [tu pti klavse~]

Further experimental data fail to show any trace of lip rounding attributable to schwa. Data was collected from one speaker who was recorded and videotaped reading six pairs of sentences, with 20 repetitions of each sentence. The pair in (1)-(2), the only one used by Barnes & Kavitskaya, was used again, as well as true minimal pairs such as the one in (3)-(4), in which rounded segments were systematically avoided. The internal vertical aperture between the lips, internal width and external width were measured in the vicinity of the clitic consonant and the corresponding word-initial one. Results show no difference in lip position between the two sentences in each pair, except for the pair in (1)-(2). The difference, however, appears to be explained by a significant difference in opening between the [v] in (2) and the [f] in (1). Barnes & Kavitskaya's result, then, is argued to follow from factors that are independent from schwa or its associated boundary.

(3) *Élie t'rame bien ça* "Élie rows that well for you" [eli t#ram bje~sa]

(4) *Élie trame bien ça* "Élie plots that well" [eli tram bje~ sa]

II. Clitic consonants before an omitted schwa, for example [d] in (5), tend to be longer and more strongly articulated than identical consonants not involving schwa (6) (Lebel 1968, Rialland 1986, Fougeron & Steriade 1997, 1999). Fougeron & Steriade (1997, 1999) propose that the clitic in (5) has the articulatory properties of a prevocalic [d] by phonetic analogy with the underlying form /dœ/.

(5) *pas d'rôle* "acting game" [pa#d#rol]

(6) *pas drôle* "funny game" [pa#d̲rol]

The relative strength of clitic consonants appears to be better understood in terms of their morphemic status rather than their underlying prevocalicness. Studies have shown that morphemic consonants are regularly longer than corresponding nonmorphemic ones (Walsh & Parker 1983). This also applies to clitic consonants in French, and explains the additional result that the clitic [d] is longer when schwa is omitted, as in (5), than when it is pronounced (*pas de rôle* [pa#dœ#rol]), a difference that underlying prevocalicness does not predict.

References

Barnes, J. & D. Kavitskaya. 2002. Phonetic analogy and schwa deletion in French. Proceedings of the Twenty-Seventh Annual Meeting of the Berkeley Linguistic Society, 39-50.

Côté, M.-H. 2000. Consonant cluster phonotactics: a perceptual approach. Ph.D. dissertation, MIT.

Fougeron, C. & D. Steriade. 1997. Does the deletion of French schwa lead to the neutralization of lexical distinctions? In Proceedings of the Eurospeech conference, Rhodes.

Fougeron, C. & D. Steriade. 1999. Au-delà de la syllabe: le rôle des informations articulatoires stockées dans le lexique pour l'analyse de la chute de schwa. In S. Wauquier-Gravelines (ed.) Journées d'études linguistiques: la syllabe sous tous ses aspects, Université de Nantes.

Lebel, J.-G. 1968. Allongement compensatoire de quelques consonnes par suite du "e" caduc. Revue de phonétique appliquée 7: 53-78.

Rialland, A. 1986. Schwa et syllabes en français. In E. Sezer & L. Wetzels (eds.) Studies in compensatory lengthening. Dordrecht: Foris.

Tranel, B. 1981. Concreteness in generative phonology: evidence from French. Berkeley: University of California Press.

Walsh, T. & F. Parker. 1983. The duration of morphemic and nonmorphemic /s/ in English. Journal of phonetics 11: 201-206.

Co-articulation and Speech Acoustics in Phonological Change: some experimental data

Paul M. De Decker
New York University

Why should some sounds change while others remain stable? Using Boersma's articulatory model, this report examines the role of articulatory variation and nasalization in the progress of an apparent sound change in Canadian English in which two variants of /ae/ exhibit different patterns. The results suggest that both articulatory variation and its relation to speech acoustics are significant factors in the progress of this diachronic change.

Clarke et al. (1995) suggest that vocalic changes in several urban centers of Ontario, Canada are the result of a chain shift in currently underway in Canadian English. Of particular interest to this study is the low-front vowel /ae/ (found in words such as *hat* and *cat),* which is reported to sound lower and more retracted in younger Canadians. Data from acoustic analyses of first and second formant frequencies (De Decker 2002) show two distinct groups within the acoustic range for /ae/. Further analyses indicate that these two sets are conditioned by the following consonant. Specifically, post-vocalic nasals and voiced velar obstruents produce both lower f1 and f2 values compared to following stops, fricatives and liquids. Examinations of the f1 and f2 using apparent time data from older and younger speakers show that f2 of younger speakers are lower than values for older speakers, suggesting a change in the phonetic quality of /ae/. However, retraction was not found for all variants of /ae/ – those preceding nasals and voiced velar stops exhibit no change between generations. Why should /ae/ sometimes be resistant to shifting while other times not?

This study attempts to answer this question. It was hypothesized that these distinct patterns for /ae/ are due to a lack of variation in nasalized variants of /ae/ compared to non-nasalized variants. This hypothesis was tested using Praat's articulatory synthesis module to model these phonetic patterns. Articulatory variation of /ae/ was simulated by producing 100 different vocal tract configurations, all of which involved adjustments made to the tongue body in two ways: the degree of constriction at the phayrngeal cavity and the relative height of the tongue (see also Goldstein 1983 for a similar experiment). These 100 tract shapes were coupled with model's nasal cavity. The acoustic result of this coupling on articulatory variation was then examined. The results show that nasalized and non-nasalized /ae/ exhibit wide distributional ranges resulting from pertubations made to the shape of the vocal tract. In particular, despite uniform articulatory changes, nasalized variants showed a marked reduction in acoustic variation. In comparison to the non-nasalized vowels, which exhibited a considerable range of variation for f1 and a wider range for f2, nasalized vowels showed only a wide distribution for f2 and a notably reduced range of variation for f1 values.

Thus, nasalization reduces the effect of articulatory variation and as a result, variation of nasalized /ae/ is restricted. Assuming an Ohalian view of sound change (Ohala 1981), it is clear that no change can result if there is no significant variation in the pronunciation of /ae/.

One unexpected pattern is the wide range of variation for nasalized /ae/. It is hypothesized that functional considerations (i.e. the maintenance of constrasts between /ae/ and neighboring vowels) would operate in reducing rampant variation. Therefore, a theory of change must consider both articulatory and acoustic factors involved in production and their perception by the listener.

References

Boersma (1998). Functional Phonology: Formalizing the interactions between articulatory and perceptual drives. Netherlands Graduate School of Linguistics.

Boersma, P. and Weenik, D. 2003. Praat. http://www.fon.hum.uva.nl/paul/ .

Clarke, S., Elms, F., Youssef, A. 1995. The Third Dialect of English: Some Canadian Evidence. Language Variation and Change. 7, 2, 209-228.

De Decker, P. 2002. Hangin' & Retractin': Adolescent Social Practice and Phonetic Variation. University of Pennsylvania Working Paper in Linguistics. 8.3.

Goldstein, L. 1983. Vowel shifts and articulatory-acoustic relations; in Cohen, van den Broecke, Proc. of the 10th Int. Congr. Phonet. Sci., pp. 267-273 (Foris, Dordrecht 1983).

Krakow, Rena A., Beddor, P.S., Goldstein, L., Fowler, C. 1988. Co-articulatory Influences on the Perceived Height of Nasal Vowels. Journal of the Acoustical Society of America, 83, 3. 1146-1158.

Ohala, J.J. (1981). The listener as a source of sound change. In C.S. Masek, R.A. Hendrick, &
M.F. Miller (Eds.), Papers from the Parasession on language and behavior (pp. 178-203). Chicago: Chicago Linguistic Society.

The effect of phonetic factors in the perception of phonological variation: An experimental study of internal constraints of syllable-final /s/

Manuel Díaz-Campos & Richard J File-Muriel
Indiana University

Previous investigations analyzing phonological variation in Spanish have pointed out that phonetic context is an important factor for predicting different variants of /s/. In many studies (Cedergren 1973, Poplack 1979, 1980, 1986, Ranson 1993, Samper Padilla 1990, Terrell 1977, 1978, 1979, 1986, among others), it is claimed that /s/ aspiration is more likely to happen word-internally, preceding a consonant, while elision is more likely to happen word-finally preceding a pause. The methodology used in those studies relies on grammatical judgments for performing the analysis of phonetic context, which could present a perceptual bias that could misrepresent the reality of the realization of the segments. Studies in the area of perception (House and Fairbanks 1953, Paterson and Lehiste 1960, Repp and Williams 1985, among others) have pointed out that in the case of voiced and voiceless consonants acoustic cues relevant to the perception of these consonants include duration of the preceding vowel as well as the properties of the stop consonant such as the release and the duration and voicing of the closure interval. To our knowledge, there have been no studies that verify whether or not these perceptual judgments indeed represent the phonetic reality of /s/. In the present investigation, we analyze patterns of perception in a variety of phonetic contexts using synthesized speech in order to determine potential differences between perception and phonetic realization in the production of the variants of syllable-final /s/.

The study focuses on the phonetic cues that might affect the perception of the variants of syllable-final /s/. Synthesized stimuli were created using MBROLA speech synthesizer and duration of the preceding vowel was manipulated in 20-ms steps. The following context was manipulated in order to observe if there is an effect in the perception of the variants of syllable-final /s/ according to phonetic environment. In general terms, we have included the following possible environments: (1) vowel, (2) pause, and (3) consonant. Regarding consonants, we have controlled for manner of articulation (i.e. stops, affricates, fricatives, nasals, and liquids) as well as place of articulation (labial, coronal, palatal, and dorsal). We have also included a classification of the following vowel according to height, backness, and roundness.

The findings of this research reveal that the length of the preceding vowel and the following consonant are important cues for the perception of the variants of syllable-final /s/. When the preceding vowel had a shorter duration, perception was more accurate, whereas when the preceding vowel was longer in duration, the number of incorrect judgments increased. Subjects had more difficulties perceiving aspiration and deletion than retention cases, which is an indication of the problems that may arise when the analysis is performed by ear. The following environments where subjects had more difficulty perceiving aspiration and deletion were continuant obstruents and pause. In the case of following continuant obstruents, subjects used the following context as phonetic cue regardless of the real phonetic realization in the target token. Regarding a following pause for aspirating and deleting targets, the number of mismatches increases due to the lack of a following segment that could be used as phonetic cue. Findings suggest that the perception of aspiration and deletion in previous research may not reflect phonetic reality. The claim that the deletion of syllable final /s/ is more likely to occur after a

pause might be purely perceptual since it is precisely in this context where subjects are less likely to distinguish between deletion and aspiration.

References

CEDERGREN, HENRIETTA. 1973. The interplay of social and linguistic factors in Panamá, Cornell University: Doctoral thesis.

CEDERGREN, HENRIETTA, NÚÑEZ, RAFAEL, PÁEZ URDANETA, IRASET and GUITART, JORGE M. 1986. Estudios sobre la fonología del español del Caribe. Caracas: Ediciones La Casa de Bello.

POPLACK, SHANA. 1980a. Deletion and disambiguation in Puerto Rican Spanish. Language, 56.371-85.

—. 1980b. The notation of the plural in Puerto Rican Spanish: Competing contraints on /s/ deletion. Locating Language in Time and Space, ed. by William Labov and David Sankoff, 55-67. New York: Academic.

RANSON, DIANA L. 1991. Person marking in the wake of /s/ deletion in Andalusian Spanish. Language Variation and Change, 3.133-52.

—. 1992. Nominal number marking in Andalsuian Spanish in the wake of /s/. Hispanic Linguistics, 4.301-27.

—. 1993. The interaction of linguistic and contextual number markers in Andalusian Spanish. Hispania, 76.919-30.

REPP, BRUNO AND DAVID WILLIAMS. 1985. Influence of following context on perception of the voiced-voiceless distinction in syllable-final stop consonants. Journal of the acoustic society of America 78.445-457.

SAMPER PADILLA, JOSÉ ANTONIO. 1990. Estudio sociolingüístico del español de Las Palmas de Gran Canaria. Santa Cruz de Tenerife: La Caja de Canarias.

TERRELL, TRACY D. 1975. Aspiration in Cuban Spanish: Theoretical observations. Revista de lingüística teórica y aplicada, 13.93-107.

—. 1977a. Constraints on the aspiration and deletion of final /s/ in Cuban and Puerto Rican Spanish. The Bilingual Review/La revista bilingüe, 4.35-51.

—. 1977b. Universal constraints on variably deleted final consonants: Evidence from Spanish. The Canadian Journal of Linguistics/La Revue Canadienne de Linguistique, 22.156-68.

—. 1978. Sobre la aspiración y elisión de /s/ implosiva y final en el español de Puerto Rico. Nueva Revista de Filología Hispánica, 27.24-38.

—. 1979. Final /s/ in Cuban Spanish. Hispania, 62.599-612.

—. 1982. Current trends in the investigation of Cuban and Puerto Rican phonology. Spanish in the United States: Sociolinguistic Aspects, ed. by Jon Amastae and Lucía Elías-Olivares.

—. 1983. Epenthesis in Spanish: A concrete analysis. Spanish in Social Context, ed. by John J. Bergen and Garland D. Bills, 67-75. Washington DC: Georgetown University Press.

The phonetic motivation of stop assibilation

T. A. Hall†, Silke Hamann* & Marzena Zygis*
*University of Leipzig† and ZAS**

The present study is concerned with stop assibilation — a process whereby stops become sibilant affricates or sibilant fricatives before high vocoids. Two examples are presented in (1a, b) from Finnish and Korean respectively. Similar examples of assibilations can be found in Romanian, Cheyene, Efik, Japanese and Quebec French (see Bhat 1978 and Kim 2001).

(1) a. $t \rightarrow s$ / __ i b. $t\ t^h \rightarrow ts\ ts^h$ / __ i

Stop assibilations are defined here as processes with the following four properties (see also Clements 1999 and Kim 2001): (a) the input segments are stops, which are usually alveolar or dental, (b) the trigger is (typically) some subset of the high front vocoids (e.g. /i y j/), (c) the output is always a sibilant (either an affricate or a fricative) and (d) the trigger is always to the right of the target.

Kim (2001) offers a phonetic explanation for these properties: The creation of sibilants from stops has its phonetic origin in the brief period of turbulence which occurs at the release of a stop into a high vocoid.

In the present study we present phonetic evidence supporting the two implications in (2), neither of which is discussed by Clements (1999) or Kim (2001):

(2) a. Assibilation of /t/ in /tj/ implies assibilation of /t/ in /ti/
b. Assibilation of /d/ implies the assibilation of /t/

Both (2a) and (2b) can be confirmed by examining the cross-linguistic evidence for stop assibilations. For example, there are languages like Quebec French and Plains Cree, in which /t/ assibilates before /i/ and /j/ and Latin, in which /t/ assibilates only before /j/ but not before /i/. No language assibilates /t/ only before /i/ but not before /j/.

In our talk we will present the results of an acoustic study of /ti tj di dj/ sequences in German (a Germanic language with assibilation) and Arabic (a non-Germanic language without assibilation). In our study we measured the frication phase after the /t d/ release (until the onset of a following vocoid) for three speakers of these two languages in each of these four sequences. We found that the friction phase for /ti/ was significantly longer than that of /tj/, supporting the implication in (2a). Furthermore, we found that the friction phase of /t/ is significantly greater than /d/, thereby lending support to the implication in (2b).

Our results are significant because they show that the universals in (2) (which are not discussed in the literature cited above on stop assibilations) can be confirmed by phonetics. By contrast, there is no clear phonological reason for why (2a, b) should hold. Thus, our study provides an example of a phonological process whose explanation can only be made clear by appealing to phonetics.

References

Bhat, D. N. S. 1978. A general study of palatalisation. In: J. Greenberg et al. (eds.) Universals of human language. 47-92. Stanford: Stanford University Press.

Clements, G. N. 1999. Affricates as noncontoured stops. In: O. Fujimura et al. (eds.) Item, order in language and speech. Prague: Charles University Press.

Kim, H. 2001. A phonetically based account of phonological stop assibilation. Phonology 18: 81-108.

From phonetic differences to phonological asymmetries: Secondary articulation contrasts in liquids

Alexei Kochetov
Simon Fraser University and Haskins Laboratories

It has been noted that secondary articulation contrasts in laterals and rhotics exhibit certain phonological asymmetries [1, 2, 3, 4, 5]. First, if a language has palatalized (or palatal) liquids, most likely these are laterals rather than rhotics. Second, palatalized laterals tend to occur in a wider range of phonotactic contexts than palatalized rhotics do. Third, front/back contrasts in laterals are relatively stable diachronically, while similar contrasts in trills/taps exhibit a strong tendency to positional or total neutralization. While previous works proposed formal synchronic accounts of these markedness asymmetries [3, 4], the role of phonetic factors in their evolution has not been thoroughly investigated. In this study we examine articulatory and acoustic properties of Russian liquids: velarized vs. palatalized dental/alveolar laterals (/l/ vs. /lʲ/) and plain vs. palatalized (post-)alveolar trills (/r/ vs. /rʲ/). The goal is to determine phonetic differences between the non-palatalized and palatalized counterparts, as well as the relation between these differences and the phonological asymmetries.

An *articulatory analysis* was based on tongue movement data from a magnetic articulometer corpus of Russian [6]. The examined data involved the consonants /l lʲ r rʲ/ produced by 3 speakers in word-final position in nonsense words (ta_#api, 5 repetitions), as well as in additional real word tokens. The results showed that the main difference between non-palatalized and palatalized consonants was in the high vs. low position of the tongue body (TB) gesture; yet, this difference was significantly larger for laterals than for rhotics (about 10 mm vs. 4 mm on average). The larger difference for laterals was in part due to tongue body lowering for the velarized /l/ (and tongue dorsum backing; [cf. 7, 8, 9]). The smaller difference for rhotics likely resulted from the substantial tongue tip (TT) raising for the plain alveolar trill, which led to the intermediate raising and backing of TB. In addition, the consonants showed some differences in primary articulation: for all speakers TT was significantly higher and more back for /lʲ/ compared to /l/; it was lower and more front for /rʲ/ compared to /r/. These differences were somewhat larger for laterals than for rhotics (about 5 mm vs. 2 mm for TT raising). Temporally, the results showed that TB was achieved and released simultaneously with TT for the palatalized lateral, while TB followed TT for the palatalized rhotic. This asynchrony in the articulation of /rʲ/ can be attributed to the conflicting demands imposed on the tongue by apical and palatal articulations [10], as well as by aerodynamics of trill production [11, 12]. An *acoustic analysis* of the articulatory data revealed that, not unexpectedly, formant differences for laterals were quite substantial during VC/CV transitions and relatively stable during their closure periods (F2 of up to 800 Hz). The formant differences between rhotics were significant, but smaller overall and less noticeable during the VC transition (F2 of up to 450 Hz). Unlike the plain trill, its palatalized counterpart had no more than one period of vibration followed by some high frequency noise (apparent effects of TB raising). In general, the results were consistent with previous x-ray/palatography-based and acoustic descriptions of Russian liquids [13, 14, 15].

In sum, the study has identified substantial articulatory and acoustic (as well as aerodynamic) asymmetries between Russian lateral and rhotic contrasts. These asymmetries are rooted in the physical properties of the consonants, and thus are, to a large degree, language-independent: lateral articulations in general allow for more flexibility in particular gestural and

acoustic realizations of secondary articulation contrasts than dental/alveolar rhotic articulations do. The observed low-level asymmetries, mediated by learner's misperceptions [16, 17; cf. 10], can be seen as the main driving force behind the evolution of the phonological asymmetries between palatalized laterals and rhotics.

References

[1] Bhat, D.N.S. 1978. A general study of palatalization. In: Universals of language, Vol. 2: Phonology, ed. J. H. Greenberg. Stanford, CA: Stanford University Press. 47-91.

[2] Maddieson, Ian. 1984. Patterns of sounds. Cambridge: Cambridge University Press.

[3] Walsh Dickey Laura. 1997. The phonology of liquids. Ph.D. dissertation, University of Massachusetts.

[4] Hall, T. A. 2000. Typological generalizations concerning secondary palatalization. Lingua 111. 1-25.

[5] Carlton, T.R. 1990. Introduction to the phonological history of the Slavic languages. Columbus, OH: Slavica.

[6] Kochetov, Alexei and Louis Goldstein 2002. A corpus of Russian magnetic articulometer data. Haskins Laboratories. New Haven, CT.

[7] Recasens, Daniel. 1999. Lingual coarticulation. In: William J. Hardcastle and Nigel Hewlett (eds.), Coarticulation: theory, data, and techniques. Cambridge, UK: Cambridge University Press: 80–104.

[8] Browman, Catherine P. and Louis Goldstein. 1995. Gestural syllable position effects in American English. In: Fredericka Bell-Berti and Lawrence J. Raphael (eds.), Producing speech: Contemporary issues (for Katherine Safford Harris). Woodbury, N.Y.: AIP Press: 19–33.

[9] Gick, Bryan, Fiona Campbell, Sunyoung Oh and Linda Tamburri-Watt. Submitted. Towards universals in the gestural organization of syllables: A cross-linguistic study of liquids. University of British Columbia.

[10] Kochetov Alexei. 2002. Production, perception, and emergent phonotactic patterns: A case of contrastive palatalization. New York: Routledge.

[11] Ladefoged, Peter and Ian Maddieson. 1996. The Sounds of the world's languages. Cambridge, M.A.: Blackwell.

[12] Kavitskaya, Darya. 1997. Aerodynamic constraints on the production of palatalized trills: The case of the Slavic trilled [r]. In: Proceedings from the 5th European Conference on Speech Communication and Technology. Volume 2. University of Patras, Rhodes, Greece. 751-754.

[13] Skalozub, Larisa G. 1963. Palatogrammy i rentgenogrammy soglasnykh fonem russkogo literaturnogo iazyka. Kiev: Izdatel'stvo Kievskogo universiteta.

[14] Bolla, K. 1981. A conspectus of Russian speech sounds. Cologne: Bölau.

[15] Zinder, Lev R., Lija V. Bondarko, and L. A. Verbitskaja. 1964. Akusticeskaja xaracteristika razlicija tverdyx i mjagkix soglasnyx v russkom jazyke. In: Voprosy fonetiki (Uchenye Zapiski LGU, No. 325), ed. by M. I. Matusevich. Leningrad State University, Leningrad. 28-36.

[16] Ohala, John. 1981. The listener as a source of sound change. In: Papers from Chicago Linguistic Society parasession on language and behavior, C.S. Masek, R.A. Hendrick, and M.F. Miller. Chicago, IL: CLS. 178-203.

[17] Ohala, John. 1990. The phonetics and phonology of aspects of assimilation. In: Papers in laboratory phonology I: Between the grammar and physics of speech, eds. J. Kingston and M.E. Beckman. Cambridge: Cambridge University Press.

Sustaining phonemic distinctions: A study of phonetic cues in Japanese long vowels

Tomoko Kozasa
University of Hawaii at Manoa

In 1946, Martinet (1955:22) states that "It is an acknowledged fact that a good many and probably most sound changes seem to be due to an insufficient effort on the part of the speakers to distinguish carefully between neighboring sounds." What then do speakers subconsciously do if they want to retain the between two phonemes? This paper examines how native speakers of the Tokyo dialect of Japanese use acoustic cues to maintain vowel length distinction.

Japanese is a pitch-accent language. An accented word has a high-low (HL) pitch sequence in the Tokyo dialect. When a word contains an accented long vowel, the HL sequence falls within the vowel (McCawley 1968). In addition to being pitch-accented, Japanese is a mora-timed language; that is, the mora is a timing unit used to determine the length of a segment. Many acoustic investigations have focused on the durational aspect of the Japanese mora (Bloch 1950, Han 1962, 1994, Homma 1981, Beckman 1982a, Port *et al.* 1987, Han 1994, Warner and Arai 2001). Only a few researchers have investigated the interaction of pitch-accent and vowel duration. On one hand, Beckman (1982b) suggests that accent has only a minimal effect on vowel duration in Japanese, and Larish (1989) reports similar findings from a production experiment involving Japanese long vowels. On the other hand, some perception experiments suggest that pitch may influence the judgment of the vowel length for native speakers of Tokyo dialect (Nagano-Madsen 1990, Kozasa 2002).

Two production experiments were conducted. One involved both duration and pitch. The other involved only duration. The materials used in the first experiment were 50 minimal pairs of accented vowels, such as *biru* (HL) 'building' and *biiru* (HLL) 'beer.' The second experiment used 20 minimal or near minimal pairs of unaccented words, such as *tokai* (HHH) 'cosmopolitan' *tookai* (HHHH) 'eastern sea' and 10 words containing long vowels without accent, such as and *meekappu* (HHHHL) 'make-up.' There were four native speakers of Tokyo dialect (two females and two males) in each experiment. They read the words in frame sentences at three different speech rates – fast, normal, and slow. Only the data from fast and slow speech were used for the present study.

Measurements of pitch fall and of short and long vowel durations showed that speech rates did not have a significant effect on pitch fall in either short or long accented vowels. But speech rates were found to affect the duration of all types of vowels ([$p < .0001$]). The ratio of mean duration between short and long accented vowels was 1:1.78 in fast speech and 1:1.91 in slow speech. For unaccented vowels, the mean duration ratio of short to long vowels was 1:1.93 in fast speech and 1:2.43 in slow speech. These results suggest that the speakers rely more on pitch fall than on duration when they produce accented long vowels, as the durational cue is not as stable. Slow speech produced more lengthening of unaccented long vowels. For these unaccented vowels, where speakers must rely solely on the durational cue, they increase the duration of the unaccented vowels to ensure that the long vowels can be heard as long. Thus, speakers' subconscious efforts may result in preserving the phonemic short-long distinction.

References

Beckman, Mary E., 1982a. Segment duration and the 'Mora' in Japanese. Phonetica 39, 113-135.

Beckman, Mary E., 1982b. Effects of accent on vowel duration in Japanese. The Journal of the Acoustical Society of America, 71, S23.

Bloch, B., 1950. Studies in colloquial Japanese IV: Phonemics. Language, 26, 86-125.

Han, M.S., 1962. The feature of duration in Japanese. Onsei no Kenkyuu, 10, 65-75.

Han, M.S., 1994. Acoustic manifestations of mora timing in Japanese. The Journal of the Acoustical Society of America, 96, 73-83.

Homma, Y., 1981. Durational relationship between Japanese stops and vowels. Journal of Phonetics, 9, 273-381.

Kozasa, Tomoko, 2002. Duration and F0 cues for vowel length in Japanese. The Journal of the Acoustical Society of America, 111, 2365.

Larish, Michael D., 1989. Converging sources of evidence again the phonetic reality of the mora in Japanese. In University of Hawai'i Working Papers in Linguistics, 19 (2), 103-151.

Martinet, A. 1955. Phonology as functional phonetics: Three lectures delivered before the University of London in 1946. Oxford, UK: Basil Blackwell.

McCawley, James D., 1968. The phonological component of a grammar of Japanese. The Hague: Mouton.

Nagano-Madsen, Yasuko, 1992. Influence of fundamental frequency pattern on the perception of the vowel mora in Japanese. Phonum, 1, 31-35.

Port, R.F., Dalby, J., and O'Dell, M., 1987. Evidence for mora timing in Japanese. In The Journal of the Acoustical Society of America, 81, 1574-1585.

Warner, Natasha and Takayuki Arai, 2001. The role of the mora in the timing of spontaneous Japanese speech. In The Journal of the Acoustical Society of America, 109, 1144-1156.

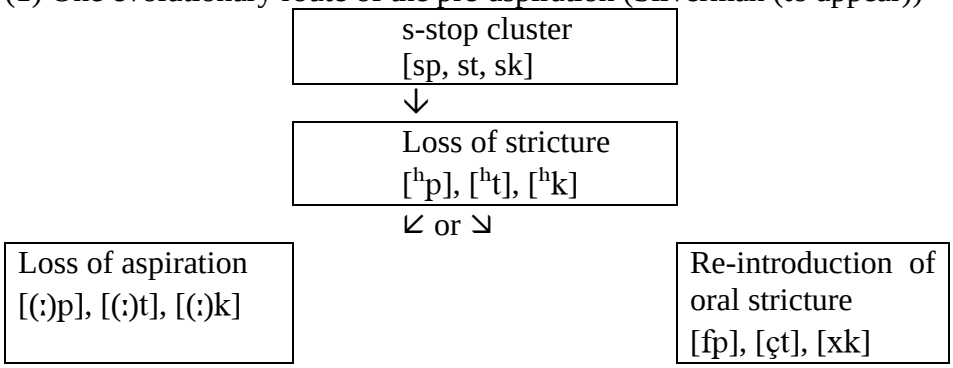
Asymmetry in Laryngeal Metathesis

Tae-Jin Yoon
University of Illinois at Urbana-Champaign

This paper addresses the main determinant that governs laryngeal metathesis, a process that conditions a reordering of segments such as /ahpa/ → [apha] in some languages. If we consider each segment as an autonomous entity then *a priori* we should expect metathesis to also produce reordering in the opposite direction, e.g., /apha/ → [ahpa]. In this paper I present evidence from numerous languages that laryngeal methathesis applies asymmetrically to produce output sequences such as [apha], and largely avoids output sequences like [ahpa], and I propose an analysis of this aysymmetry in terms of perceptual factors. Superficially, the only difference between the two reorderings, *ahpa* and *apha*, is the linear sequence of the two segments *h* and *p*. Yet it is also the case that this sequencing difference results in aspiration preceding a stop in *ahpa*, and in aspiration following stop release in *apha*. Drawing on the work of Silverman (to appear), I argue that the two sequences do not have equal status in phonological systems cross-linguistically. I show that the observed asymmetry in laryngeal metathesis has a plausible basis in perceptual factors, and offer new evidence from laryngeal metathesis in support of Hume's (to appear) perceptually-based model of metathesis.

In a typological study of pre-aspirated stops, Silverman (to appear) concludes that pure cases of pre-aspiration are extremely rare. Pre-aspirates, many instances of which derive from original s-stop clusters, are affected by processes of sound change typically resulting in the loss of aspiration or enhancement of aspiration through fricative noise, as illustrated in the evolutionary sound changes shown in diagram (1).

(1) One evolutionary route of the pre-aspiration (Silverman (to appear))



Silverman views the loss or enhancement of the pre-aspiration component as grammatical responses to the weak aerodynamic/acoustic/auditory cues of the pre-aspirate -- a prime example of perceptually-conditioned sound change. I show that in many languages laryngeal metathesis provides yet another strategy for resolving pre-aspirate stop sequences. Examples of metathesis in synchronic and diachronic sound changes are aadduced in languages such as Icelandic (Ringen 1999), Gujarati (Cardona 1965), Cherokee (Flemming 1996), Yuman (Langdon 1976), Mandaic (Malone 1971), Harari (Leslau 1963, Hume (to appear)), and Kiliwa (Mixco 1971), among other languages. In these languages, the laryngeal component of a consonant or consonant cluster undergoes metathesis, relocating to a position where the perceptual cue to aspiration is more salient. I observe two types of asymmetry in laryngeal metathesis involving

aspiration. First, pre-aspirate plus stop sequences undergo metathesis resulting in stop plus post-aspirate, but the reverse pattern is not observed. Second, the reordering resulting from metathesis varies for different input consonants according to the perceptual salience defined by the local phonological context. When a language exhibits laryngeal metathesis affecting both stops and non-stop consonants in sequence with aspiration, metathesis locates the laryngeal component to the site preceding a fricative or non-vocalic sonorant, and to the site following a stop consonant.

This work contributes to Silverman's analysis of the rarity of pre-aspirates through the addition of many new examples exhibiting laryngeal metathesis, and supports the more general claim that sound changes, including those affecting pre-aspirates, are in many instances perceptually motivated (Steriade 2001). The analysis of laryngeal metathesis put forth in this paper also provides important evidence for Hume's (to appear) Ambiguity/Attestation Model of metathesis. Hume addresses several different properties of metathesis cross-linguistically, and states that the order of segments resulting from metathesis seems to have, more often than not, better acoustic/auditory cues than the unmetathesized sequence. Hume claims that the patterns of metathesis can be explained when "the phonetic nature of the sounds involved as well as the speaker/hearer's knowledge of native sound patterns and their frequency of occurrence" are considered. While Hume presents many examples of synchronic metathesis in support of her proposal, regarding laryngeal metathesis Hume states that the direction of change can not be predicted in terms of acoustic/auditory cues alone, but consideration must also be given to the sound patterns of the language. The examples of both synchronic and diachronic metathesis affecting pre-aspirates presented in this paper show that perceptual salience is in fact the main determinant of the output order resulting from metathesis for those cases where there would be a difference between possible output orders in terms of perceptibility of the aspiration component. The aspiration component moves to a position in the surrounding phonological structure in which the perceptual cues to aspiration are more salient than in the original position. Thus, the variety of outcomes is itself predicted on perceptual grounds and laryngeal metathesis involving pre-aspirates, in general, provide important confirmation of Hume's model.

Selected References

Blevins, J. & A. Garrett (to appear). *The Evolution of Metathesis*. to appear in a volume edited by B. Hayes, R. Kirchner & D. Steriade.

Hume, E. 2001. Metathesis: Formal and Functional Considerations. in Hume, E., N. Smith and J. van de Weijer (eds.) *Surface Syllable Structure and Segment Sequencing*. HIL Occasional Papers. Leiden, NL: HIL.

Hume, E. (to appear). Predicting Metathesis: The Ambiguity/Attestation Model. *Language* 80(2)
[Downloaded from http://www.ling.ohio-state.edu/~hume/metathesis]

Silverman, D. 1997. *Phasing and recoverability*. Outstanding dissertations in linguistics series. New York: Garland.

Silverman, D. (to appear). On the rarity of pre-aspirated stops. *Journal of Linguistics.*

Steriade, D. (2001). The Phonology of Perceptibility Effects: the P-map and its consequences for constraint organization. ms. MIT.
[Downloaded from http://web.mit.edu/linguistics/www/bibliography/steriade.html]

Vowel duration in Kinyarwanda: Effects of quantity, height and position

Scott Myers
University of Texas at Austin

In Kinyarwanda, a Bantu language of Rwanda, there is a contrast in vowel length as evidenced by such pairs as *gutaka* "to scream" and *gutaaka* "to decorate". The contrast is neutralized in various positions (Sibomana 1974, Kimenyi 1979):

(a) Word-initially or word-finally, only short vowels occur.

(b) Before a nasal – obstruent sequence (**pre-NC**), only long vowels occur.

(c) After a consonant – glide sequence (**post-CG**), only long vowels occur.

All else being equal a long vowel has a greater duration than a corresponding short vowel. But many factors have been shown to affect the duration of vowels. For example, vowels that are lower in the vowel space are longer than those that are higher in that space (Lehiste 1970), stressed ones are longer than unstressed ones (Lehiste 1970, Beckman et al. 1992), phrase-final ones are longer than non-final ones (Oller 1972, Klatt 1976, Wightman et al. 1992), and vowels at a slower rate are longer than corresponding vowels at faster rates (Lehiste 1970, Magen and Blumstein 1993). The result is that a short vowel in one condition can actually have a greater duration than a long vowel in another.

To investigate how such factors affected the durational contrast between long and short vowels in Kinyarwanda, four adult speakers of the language were recorded producing short phrases of 6 syllables. The test vowels varied with respect to the following categories:

(a) *Vowel type*: long, short, pre-NC, post-CG

(b) *Vowel height*: high, low

(c) *Phrase position*: word-initial/phrase-initial, word-penultimate/phrase-medial, word-final/phrase-medial, word-initial/phrase-medial, word-penultimate/phrase-penultimate, word-final/phrase-final

Each speaker produced 24 tokens in each condition – half at a self-selected normal rate and half as fast as they could.

As expected, long vowels had significantly longer duration than short vowels, and low vowels were longer than high vowels. Vowels in the pre-NC context were intermediate in duration between short and long vowels, as has been found for some other Bantu languages (Maddieson and Ladefoged 1993, Hubbard 1995).

Phrase-penultimate short vowels were significantly longer than short vowels in any other position, and vowels in word-penultimate position were the second longest class. This suggests that the penultimate position is a prominent, stressed position.

The phrase-final vowel was subject to complete or near-complete devoicing. The duration of the devoiced vowel varied dramatically among the speakers: for some it was shorter than in any other position, while for others it was longer than any other unstressed position.

These findings bear on both the phonological representation and the phonetic implementation of phonemic vowel length.

References

Beckman, M., Edwards, J. and Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. In G. Docherty and D. Ladd (Eds.) Papers in laboratory phonology II: Gesture, segment, prosody (pp. 68-86). Cambridge: Cambridge University Press.

Bickmore, L. (1989). Kinyambo prosody. PhD dissertation, UCLA, Los Angeles.

Hubbard, K. (1995). "Prenasalised consonants" and syllabic timing: Evidence from Runyambo and Luganda. Phonology 12, 235-256.

Kimenyi, A. (1979). Studies in Kinyarwanda and Bantu phonology. Carbondale, Il.: Linguistic Research Inc.

Klatt, D. (1976). Linguistic uses of segment duration in English: Acoustic and perceptual evidence. JASA, 59, 1208-1221.

Lehiste, I. (1970). Suprasegmentals. Cambridge: MIT Press.

Maddieson, I. and Ladefoged, P. (1993). Phonetics of partially nasal consonants. In M. Huffman and R. Krakow (Eds.) Nasals, nasalization and the velum (Phonetics and phonology, volume 5) (pp. 251-301). Orlando, Florida: Academic Press.

Magen, H. and Blumstein, S. (1993). Effects of speaking rate on the vowel length distinction in Korean. Journal of Phonetics, 21, 387-409.

Oller, D. (1973). The effect of position in utterance on speech segment duration in English. JASA, 54, 1235-1247.

Sibomana, L. (1974). Deskriptive Tonologie des Kinyarwanda. Hamburg: Helmut Buske Verlag.

Wightman, C., Shattuck-Hufnagel, S., Ostendorf, M., and Price, P. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. JASA, 91. 1707-1717.

Flatly contradicted: The non-equivalency of a distinctive feature in Hindi and Levantine Arabic

Ryan Shosted
University of California-Berkeley

Jakobson, Fant, and Halle (1952:31) described the primary acoustic correlate of the distinctive feature [flat] as the "downward shift of a set of formants" and the primary articulatory correlate as constriction at the labial and/or pharyngeal place of articulation. They exemplify [flat] with the pharyngeal or "emphatic" consonants of Standard Arabic and the retroflex consonants found in "various languages of India" (34). During the production of the latter, they claim that the pharynx is contracted and the resonating cavity elongated, thus contributing to the downward trajectory of adjacent vowel formants. Ohala (1985) cites loanword evidence from Standard Arabic to Argobba in order to assert that languages which employ only one articulatory mechanism (like lip-rounding) to exploit the plain/flat distinction will use this native mechanism to encode words from a language that exploits a different "flattening" mechanism (in this case, pharyngealization).

In its long history of contact with Arabic, Hindi has adopted many loanwords whose source language input includes pharyngeal consonants. In accordance with Ohala's hypothesis about the functional equivalence of Argobba labialization and Arabic pharyngealization, one might expect to find the Arabic pharyngeal consonants encoded as (flat) retroflexes when adopted into Hindi. An investigation of a Hindi dictionary that lists the provenance of foreign loans, however, shows that this is not the case (McGregor 1997). Arabic pharyngeal [t?] invariably emerges as plain [t] in Hindi; [s?] as [s]; [d?] as [z]; and [th?] as [z]. In no case does Arabic [t?] ever emerge as retroflex [T] in Hindi.

To examine the perceptual motivations of this non-equivalency among flattening mechanisms, a comparative spectrographic analysis of pharyngeal and retroflex consonants was carried out. Three speakers of Levantine Arabic (LA) and three speakers of Hindi were recorded uttering nonce VCV syllables where C alternated between plain and flat stops (both voiced and voiceless), V was chosen from the set [a i u] common to both Hindi and LA, and V1 was identical to V2 for each token. The downward trajectory of the vowel formants adjacent to the consonant was quantified for each token by using scripts and averaging techniques devised for Praat. The tokens were then presented to ten LA and ten Hindi speakers. (LA was chosen because speakers preserve the pharyngeal pronunciation of the emphatics). In a forced-choice perception experiment, both sets of listeners heard a randomized list of equal numbers of plain and flat tokens spoken by the three Hindi and three LA speakers. Both Hindi and LA speakers were asked to choose the native orthographic representation for the sound they heard. Listeners were not told that they were hearing sounds from any language other than their own.

Statistical analysis of listener responses was used to determine the extent to which LA and Hindi speakers [1] perceived non-native flat consonants as their native flat counterparts and [2] judged flatness as a function of downward formant trajectory. Preliminary results for [1] indicate asymmetric behavior among Hindi and LA listeners. LA listeners appear more likely than their Hindi counterparts to categorize non-native flat consonants as native flats, while Hindi listeners show a propensity for categorizing Arabic flats as native plains. Preliminary results of the acoustic analysis [2] suggest that judgment of native flat consonants is strongly correlated with downward formant trajectories. While the results of the study are not yet final (the

responses of all 20 listeners have not yet been evaluated), the trends that appear so far call into question the functional equivalency of the feature [flat] across languages and corroborate the loanword phenomena mentioned for Hindi. It suggests that the acoustic cues for retroflexes and pharyngeals, while similar in some respects, are still dissimilar enough to engender variations in perception across speakers and languages. This presumably results in changing diachronic patterns, especially in the context of language contact and bilingualism.

Works consulted:

Al-Ani, Salman H. 1970. Arabic phonology: An acoustic and physiological investigation. Mouton: The Hague.

Bhat, D.N.S. (1974) "Retroflexion and retraction." Journal of Phonetics 2: 233-237.

Jakobson, Roman, C. Gunnar M. Fant, and Morris Halle (1952). Preliminaries to speech analysis: The distinctive features and their correlates. Technical Report 13. Massachusetts: Acoustics Laboratory, MIT.

Ladefoged, Peter and Peri Bhaskararao (1983) "Non-quantal aspects of consonant production: A study of retroflex consonants." Journal of Phonetics 11: 291-302

Leslau, W. (1957) "The phonetic treatment of the Arabic loanwords in Ethiopic." Word 13, 100 123.

McGregor, R.S. (1997). The Oxford Hindi-English Dictionary. Delhi: Oxford University Press.

Obrecht, N.S. (1968). Effect of the second formant in the perception of velarization in Lebanese Arabic. Mouton: The Hague.

Ohala, John J. (1985). "Around flat." In Victoria Fromkin, ed., Phonetic linguistics: Essays in honor of Peter Ladefoged, 223-241. Florida: Academic Press.

Distributional and statistical bases of allophonic groupings

Katherine S. White[†], Cecilia Kirk[†], James L. Morgan[†], Sharon Peperkamp*[‡] & Emmanuel Dupoux*

Brown University[†], Laboratoire de Sciences Cognitives et Psycholinguistique, EHESS - ENS – CNRS, Paris and Université de Paris 8[‡]*

One issue central to acquisition of phonological and lexical organization of language but not yet empirically addressed concerns how infants learn which segmental categories group together into phonemes. Segments in a given language do not all have the same function: some segments are phonetic variants of others, occurring only in certain contexts. For example, in Italian, intervocalic [s] is changed into [z]; [z] occurs in no other position; consequently, [s] and [z] appear in complementary distribution. Phonemic categories group basic segments together with their phonetic variants, or allophones. Thus, in Italian, the phonemic category /s/ encompasses the allophone [z]. Peperkamp & Dupoux (2002) proposed that infants acquire phonemic categories through distributional analysis of segments in the input. However, in addition to occurring in a more restricted set of contexts, variant allophones may typically occur with lesser frequency. Which (if either) of these asymmetries is important for infants' learning of allophonic groupings?

Infants were exposed to an artificial language in which certain consonantal phonemes had allophones. For example, infants in one group in which voiced and voiceless stops were allophones were familiarized with randomized sequences of "function"+"content" word pairs such as "rot pevi na bevi na suma na zuma rot tula na dula rot vanu rot fanu...", whereas infants in a second group in which voiced and voiceless fricatives were allophones were familiarized with "rot suma na zuma na pevi na bevi rot fanu na vanu rot tula rot dula..." Infants in both groups were then tested on preference for passages like:

rot pazo rot pazo na bazo rot pazo na bazo...

na zine na zine rot sine na zine rot sine...

All items were grammatical for infants in both groups. For infants in the stop allophone group, /pazo/ and /bazo/ should be one word, whereas /zine/ and /sine/ should be two words; the opposite was true for infants in the fricative allophone group. As in previous work by Best & Jones (1998) and Maye, Werker, & Gerken (2002) using a similar technique, infants listened differentially to passages containing what was perceived as one vs. two words. Across a series of experiments, the number and frequency of contexts in which allophonic variants occurred was manipulated. Cross-experiment analyses reveal the relative contributions of distributional and statistical asymmetries to infants' learning of allophonic groupings. We interpret and discuss our findings in light of current statistical learning approaches to language acquisition.

References

Best, C. T., & Jones, C. (1998) Stimulus alternation preference procedure to test infant speech discrimination. Infant Behavior and Development, 21 (Special Issue), 295.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. Cognition, 82, B101-B111.

Peperkamp, S. & E. Dupoux (2002) Coping with phonological variation in early lexical acquisition. In I. Lasser (ed.) The Process of Language Acquisition (pp. 359-385). Frankfurt: Peter Lang.

Gestural Analysis on Mandarin Tongue Twisters

Man Gao
Yale University

Recent kinematic data tracking the movements of tongue and lips during the production of repetitive sequences has found evidence for a mode-locking account of speech errors (Pouplier, 2003). The constricting organs of the vocal tract appear to function as coupled oscillators in this task, and gestural intrusion errors are observed that can be analyzed as transitions from a 2:1 mode of oscillator coupling to a more stable 1:1 mode.

This study reports an experiment on Mandarin tongue twisters that attempts to generalize the mode-locking account to non-repetitive tasks and to another language. In addition, it tests the hypothesis that a 2:1 relation among constrictors is particularly unstable (and therefore a potential tongue twister) when the constrictions are also produced *in phase*, i.e., in the syllable onset. Mandarin tongue twisters containing words/phrases with syllable onsets alternating among labial, velar and labial-velar segments were recorded, e.g., [xoŋ fɤŋ xwɑŋ] 'red phoenix', in which Lip and Tongue Dorsum gestures are in 2:1 mode with each other (in e.g., [fɤŋ xwɑŋ]). Pseudo-twisters created by replacing the labial-velars (hypothesized to be the source of the unstable 2:1 coupling) with plain velars or labials were also recorded, in the corresponding pseudo-twister of the above example, [xɑŋ] is used to replace [xwɑŋ]. In the resulting materials, Lip and Tongue Dorsum constrictions alternate with one another, but no longer are in a 2:1 relation. Mandarin speakers listened to the recordings and noted any errors. Perceived errors were also analyzed acoustically.

Results show very large differences in error rate between the real and pseudo-twisters. This supports the prediction that a 2:1 relation between gestures coordinated in phase is particularly unstable and triggers errors. In addition, as in a repetitive task, gestural intrusion errors occurred more frequently than gestural reduction ones, an outcome predicted by the mode-locking analysis (Pouplier, 2003). Language-specific issues in regard to error patterns are discussed, including the interaction between intrinsically stable modes of an oscillating system and learned lexical coordination ; and the competition between different levels of linguistic structures (morphological vs. phonetic) on patterns of gestural intrusion errors.

References:

Pouplier, M (2003). Units of Phonological Encoding: Emprical Evidence., PhD dissertation, Yale Univeristy, New Haven, CT.

The Bases of Speech Errors: Local and Non-Local Phonetic Traces

Matthew Goldrick & Sheila Blumstein
Brown University

In tongue twisters, segments from nearby syllables can intrude and cause speech errors. Research has suggested that these productions may not simply be category substitution errors but may instead reflect traces of the original target. For example, some /k/→[g] errors appear to be intermediate forms reflecting a mixture of the phonetic properties of [k] and [g] (Frisch & Wright, 2002; Lavas, 2002; Mowrey & MacKay, 1990; Pouplier et al., 2000).

The current study investigated the presence of traces in speech errors to answer two questions. The first was whether intermediate tokens found in previous work reflect the presence of traces or alternatively are a by-product of variability in phonetic realization. To investigate this question, obstruent voicing errors were analyzed in syllable initial position to determine if the distribution of speech error tokens was shifted away from that of correct targets of either category. If the distribution of /k/→[g] tokens is skewed relative to correct [g] tokens, it would suggest that these productions reflect traces of /k/—not merely the variability of [g] tokens.

Second, we investigated whether multiple acoustic cues intrude in speech errors. The implementation of a phonetic dimension is cued by both local cues (i.e., intrinsic parts of the segment) and non-local cues (i.e., part of the context in which the segment appears). Locally, syllable-initial voiceless stop consonants have: longer voice onset times (Lisker & Abramson, 1964); louder bursts (Repp, 1979); and higher F1 at vowel onset (Lisker, 1975). Voicing is also cued by acoustic properties not local to the consonant; specifically, vowels are shorter following voiceless obstruents (Peterson & Lehiste, 1960). Previous studies have focused on how local phonetic cues intrude, and have often focused on a single attribute of the consonant. None have investigated whether non-local cues will intrude as well.

Method. Four initial obstruents (/k, g, t, d/) were paired with 5 vowels (/i, E, AU, aI, Oi/) and 4 final fricatives (/f, v, s, z/). Eighty tongue twisters were then generated using a sequence of alliterating initial consonants with a constant vowel-fricative combination (e.g., "keff geff geff keff"). Five English speakers read each twister aloud 3 times quickly to induce speech errors.

Results. Productions resulting in perceptible errors (e.g., "keff→geff") were paired within each speaker with correct productions matched to the target ("keff→keff") and correct productions matched to the error ("geff→geff"). Consistent with the production of traces, the VOT distribution of error tokens was in between the distributions of correct productions matched to the target and error. For example, /k/→[g] tokens exhibited a mean VOT (32 msec) that was not only significantly shorter than correctly produced [k]s (71 msec) but also significantly longer than correctly produced [g]s (24 msec).

Although the VOT error distribution was in between those of correct productions, it was much closer to productions matched to the error (e.g., the mean VOT of /k/→[g] tokens was closer to [g] than [k]). The other local acoustic cues patterned in the same fashion. In contrast, the non-local cue, vowel length, failed to intrude in the speech errors. The mean vowel length of error tokens was no different from correct productions matched to the target. For example, the mean vowel length of /k/→[g] tokens was 139 msec, more similar to that of [k] (141 msec) rather than [g] (165 msec).

Conclusions. The distribution of speech error tokens suggests that target representations can be gradiently activated within articulatory processes. However, only local cues appear to be

activated under such conditions. This selective intrusion of local cues suggests a role for sub-syllabic units of articulatory implementation. The lack of non-local cue traces may reflect the minimal role such cues play in signaling phonetic contrasts.

References

Frisch, S. A., & Wright, R. (2002). The phonetics of phonological speech errors: An acoustic analysis of slips of the tongue. Journal of Phonetics, 30, 139-162.

Lavas, A. L. G. P. (2002, June). An acoustic study of induced speech errors in bilingual speakers. Paper presented at LabPhon 8, New Haven, CT.

Lisker, L. (1975). Is it VOT or a first-formant transition detector? Journal of the Acoustical Society of America, 57, 1547-1551.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. Word, 20, 384-422.

Mowrey, R. A., & MacKay, I. R. A. (1990). Phonological primitives: Electromyographic speech error evidence. Journal of the Acoustical Society of America, 88, 1299-1312.

Peterson, G. E., & Lehiste, I. (1960). Duration of syllable nuclei in English. Journal of the Acoustical Society of America, 32, 693-703.

Pouplier, M., Chen, L., Goldstein, L., & Byrd D. (2000, June). Articulatory evidence of gradient speech errors. Poster presented at LabPhon 7, Nijmegen, NL.

Repp, B. H. (1979). Relative amplitude of aspiration noise as a voicing cue for syllable-initial stop consonants. Language and Speech, 22, 173-189.

Ranges and levels of predictability effects on word durations in conversational English

Alan Bell*, Jason Brenier*, Michelle Gregory†, Dan Jurafsky‡ & Cynthia Girand*
University of Colorado-Boulder, State University of New York-Buffalo† and Stanford University‡*

There is ample reason from previous research to suppose that there are substantial associations between word predictability and reduced pronunciations. (See, among others, Jespersen 1923, Fidelholz 1975, Bybee 2000, Fowler and Housum 1987, Jurafsky et al. 2000.) This study concerns details about the effects of a word's predictability on its duration, specifically addressing some questions that must be answered before they can be integrated into speech production models. We focus on identifying **direct** effects not mediated by related factors also strongly associated with a word's predictability, such as the length and form of the word, whether it is accented or not, or whether it is part of an allegro or lento stretch of speech. For such associations, we further consider to what extent they are **general**, that is affecting the entire range of words in conversation in a similar way, or else affecting some classes of words differently.

The data for this study consists of about 13,000 words from the ICSI subcorpus (Greenberg 1996) of the Switchboard corpus that have been transcribed for intonational accent and phrasing (Shattuck-Hufnagel & Ostendorf 1999, Taylor 2000.) Samples of one word from each intonational phrase are drawn from the subcorpus, yielding datasets of about 1400 words. The dependent variable of the analyses is the duration of the word. Factors either controlled by selection or by variables in multiple regression analyses are articulation rate, phonological form, utterance position, fluent versus disfluent speech, intonational accent, intonational phrase position, and content/function word status. Predictability variables include lexical frequency, conditional and joint probabilities of a word with previous and following words, and measures of semantic association with the previous discourse.

For content words, lexical frequency has the strongest shortening effect on duration, the highest frequency words being about 20 percent shorter than the lowest frequency ones. This might be associated with the level of lexical access. Further questions to be addressed are the generality of this effect over common and uncommon words (the effect is stronger for higher frequency words), and to function words, which may show ceiling effects. We also find strong contextual predictability effects from preceding and following words for frequent function words (Bell et al. 2003), but **only** with the following word for content words. On the one hand, this shows that the mechanisms for the effects cannot be part of lexical access, but are likely part of phonological encoding of the prosodic domains, or part of phonetic encoding. On the other, the striking difference between frequent function words and content words raises questions such as whether the difference is due to the higher frequency of the function words, or to structural differences between content and function words. So far, although we have found no direct effects of non-local predictability from semantic relatedness, this appears to be due to weaknesses of the measures employed. An improved metric based on Latent Semantic Analysis distances of target words to automatically extracted keywords of the prior context will be tested on the ensemble of content words.

References

Bell, Alan, Daniel Jurafsky, Eric Folser-Lussier, Cynthia Girand, Michelle Gregory, and Daniel Gildea. 2003. Effect of disfluencies, predictability, and utterance position on word form variation in English conversation. J. Acoust. Soc. Am. 113.1001-1024.

Bybee, Joan. 2000. The phonology of the lexicon: Evidence from lexical diffusion. Barlow, M. and S. Kemmer, eds., Usage-based Models of Language, 65-85. CSLI, Stanford.

Fidelholtz, J. Word frequency and vowel reduction in English CLS 75.200-213.

Fowler, C. A. and J. Housum. 1987. Talkers' signaling of 'new' and 'old' words in speech and listeners' perception and use of the distinction. J. Memory and Language 26.489-504.

Greenberg, S. et al. 1996. Insights into spoken language gleaned from phonetic transcription of the Switchboard corpus. ICSLP –96, S24 –S27.

Jesperson, O. 1923. Language. Henry Holt, New York.

Jurafsky, D., A. Bell, E. Fosler-Lussier, C. Girand, and W. D. Raymond. Probabilistic relations between words: Evidence from reduction in lexical production. Bybee, J. and P. Hopper, eds., Frequency and the Emergence of Linguistic Structure, 229-254.

Shattuck-Hufnagel, S., & Ostendorf, M. 1999. Posh labeling guide—version 1.0. ms.

Taylor, P. 2000. Analysis and synthesis of intonation using the Tilt model. Journal of the Acoustical Society of America, 107 (3): 1697-1714, 2000.

Phonological Differences Emerging from Slow versus Fast Lexicon Expansion

Melissa A. Redford* & Risto Miikkulainen†
University of Oregon and University of Texas†*

Children with expressive specific language impairment (SLI-E) acquire lexical items at a much slower rate than their typically-developing peers and have persistent deficits in syntax, morphology, and phonology. The language deficits are interesting because these children only also exhibit a speech perception deficit (Tallal, et al., 1996; Wright, et al., 1997); they otherwise have normal hearing and intelligence (Johnston, 1984). Previous research into how speech perception affects grammatical development in SLI-E has been aimed primarily at the morpho-syntactic deficits (e.g., Leonard & Eyer, 1996; Joanisse & Seidenberg, 1998). A causal link between impaired speech perception and atypical phonological development is assumed, but not well-explored. Here, we consider the link in more detail and show that an additional variable—the rate of lexicon expansion—is important for understanding the different phonologies of SLI-E and typically-developing children.

We suggest that the relationship between speech perception and phonological development is mediated by the perceptual distinctiveness of the input, which is later (imperfectly) reflected in the child's productions. Impaired speech perception may lead SLI-E children to prefer and replicate only the most perceptually distinctive sounds and sound shapes of their language. Such a preference would explain their reduced phonetic inventory and simpler syllable structure (Stoel-Gammon, 1989; Rescorla & Ratner, 1996; Pharr, et al., 2000), since the most distinctive sounds are a subset of all of the sounds in an inventory and simple syllable shapes are more distinctive than complex ones (e.g., Redford & Diehl, 1999; Redford, et al., 2001).

There is, however, a problem with this hypothesis: perceptual distinctiveness is also invoked to explain normal phonological development (e.g., Kuhl, 2000). How can atypical and typical phonologies emerge under the same pressures? We addressed this question using a computational model and show that phonological differences can emerge under the same perceptual distinctiveness pressures if development occurs at different rates. In this way, we incorporate the fact that SLI-E children acquire lexical items more slowly than typically-developing children into an explanation of the phonological differences between the two groups.

Specifically, we compared the sound structure of developing artificial lexicons that reached the same final size, but achieved this size either at a slow or fast rate of expansion. Apart from the difference in expansion rate, the lexicons were developed under exactly the same conditions, i.e., under the same perceptual distinctiveness pressures and using the same underlying set of articulatorily-plausible syllables. The simulations showed that substantial structural differences emerged as a function of lexicon expansion rate. The basic structure of slow-growth lexicons changed very slowly from an initially simple structure, whereas the structure of fast-growth lexicons changed dramatically with the rapid influx of new words. These developmental differences in structure were paralleled by differences in the overall perceptual distinctiveness of the lexicons: the slow-growth lexicons maintained an optimal level of perceptual distinctiveness throughout development, whereas the fast-growth lexicons underwent a temporary, but dramatic decrease in distinctiveness with the abrupt increase in phonological complexity. Although fast-growth lexicons were allowed to stabilize and distinctiveness was regained, some structural differences persisted. Fast-growth lexicons had more word-medial

consonant sequences and more multisyllabic words than slow-growth lexicons. These differences, insofar as they suggest differences in phonological complexity, are in the direction of the observed differences between the phonologies of typically-developing children (fast-growth) and SLI-E children (slow-growth).

Overall, the simulation results suggest that the rate at which a lexicon develops has implications for its eventual structure. This observation, applied here to language acquisition, might also be relevant for understanding the different kinds of phonological changes that result from light versus heavy borrowing in the world's languages.

References

Joanisse, M., Seidenberg, M. 1998. Specific language impairment: A deficit in grammar or processing? Trends in Cognitive Sciences, 2, 240-247.

Johnston, J. 1984. Specific language disorders in the child. In N. Lass, J. Northern, L. McReynolds, D. Yoder (Eds.), Handbook of Speech-Language Pathology and Audiology (pp. 685-715). Philadelphia: B.C. Decker.

Kuhl, P.K. 2000. A new view of language acquisition. Proceeding of the National Academy of Sciences, 97, 11850-11857.

Leonard, L., Eyer, J.1996. Deficits of grammatical morphology in children with specific language impairment and their implications for notions of bootstrapping. In J. Morgan & K. Demuth (eds.), Signal to Syntax (pp. 233-248). Mahway, NJ Lawrence Erlbaum.

Pharr, A., Ratner, N. B., Rescorla, L. 2000. Syllable structure development of toddlers with expressive specific language impairment. Applied Psycholinguistics, 21, 429-449.

Redford, M., Diehl, R. 1999. The relative perceptual distinctiveness of initial and final consonants in CVC syllables. Journal of Acoustical Society of America, 106, 1555-1565.

Redford, M., Chen, C.C-, Miikkulainen, R. 2001. Constrained emergence of universals and variation in syllable systems. Language and Speech, 44, 27-56.

Rescorla, L., Ratner, N.B. 1996. Phonetic profiles of typically developing and language-delayed toddler. Journal of Speech and Hearing Research, 39, 153-165.

Stoel-Gammon, C. 1989. Prespeech and early speech development of two late talkers. First Language, 9, 207-224.

Tallal, P., Miller, S., Bedi, G., Byma, G., Wang, X., Nagarajan, S., Schreiner, C., Jenkins, W., Merzenich, M., 1996. Language comprehension in language-learning impaired children improved with acoustically modified speech. Science, 271, 81-84.

Wright, B., Lombardino, L., King, W., Puranik, C., Leonard, C., Merzenich, M. 1997. Deficits in auditory temporal and spectral resolution in language-impaired children. Nature, 387, 176-178.1997.

Asymmetries in Intermediate Stages of Cluster Acquisition

Della Chambless
University of Massachusetts

Previous studies on cluster acquisition have shown that word initial clusters do not emerge uniformly in a child's developing grammar; specifically, /*sC*/ and /*stop-liquid*/ clusters behave differently (Barlow 1997, 2001; Fikkert 1994). Little research has considered the same phenomenon in medial position. While the initial state in phonological acquisition is characterized by cluster reduction everywhere, given assumptions regarding universality of constraints and the acquisition process, intermediate stages of cluster acquisition can be predicted. In this study, I investigate whether the position (*word-initial* v. *word-medial*) and type (*stop-liquid* v. *s-stop*) of cluster affect the patterns of cluster reduction present in the speech of 2 yr. olds.

Medial position differs from initial position in that an intervocalic /CC/ sequence can be syllabified as a complex onset ($V.C_1C_2V$) or as a coda-onset sequence ($VC_1.C_2V$). If codas are present in a child's speech (*Faith >> NoCoda*), preservation of both consonants in intervocalic position is possible even if reduction ensues in word-initial position (**Complex >> Faith)*. Furthermore, *Syllable Contact* may constrain which type of cluster is permitted medially. Below I show how specific markedness constraints motivate asymmetrical cluster production, where asymmetries may refer either to cluster type (*stop-liquid* v. *s-stop*) or to position (*initial* v. *medial*).

(1) Markedness constraints which target only a subset of /CC/ sequences

Constraint	**definition**	**Role in driving asymmetrical production patterns**
*Complex	no complex onsets	When initial clusters are ruled out, medial clusters may be ok **VC.CV; *#CC**
SSP	sonority rises across the onset	i. **#br** ok; **#sp** ruled out (cluster type asymmetry) ii**. VspV** ok; **#sp** ruled out (asymmetry of position)
Syll-Con	sonority drops across syllable boundary	**Vs.pV** ok; **Vb.rV** ruled out

The differences in constraint violations incurred by the clusters above predict asymmetries in acquisition. Assuming that children follow different learning paths, some are predicted to pass through a stage in which intervocalic clusters, but not initial clusters are faithfully produced. And because *Syll-Con* favors [sp] over [br] as a heterosyllabic cluster (=medially) and *SSP* favors [br] as a complex onset, we expect an interaction between cluster type and position.

In this experiment children produced nonce words varying in type and position of cluster. ANOVA results include a significant main effect of position, $p < .001$ and a significant interaction between position and cluster type, $p<.001$. In initial position, [br] clusters are favored over [sp] clusters. In medial position, there is non-significant trend towards [sp] clusters over [br] clusters. I take the disappearance of a preference for [br] as evidence that heterosyllabic [sp] is preferred to heterosyllabic [br]. This is evidence for the role of Syllable-Contact in child acquisition.

Results of individuals show variation in cluster grammars. A partial typology emerges in which different rankings of relevant markedness constraints with Faith account for different stages of cluster production. Importantly, we observe intermediate stages in which cluster production is non-uniform across both type and position. Furthermore, differences in grammars across children is claimed to be partly a result of children following different learning paths.

Selected References

Barlow, J. (2001). The structure of /s/-sequences: Evidence from developing systems. *Journal of Child Language*, 28, 291-324.

Boersma, P. (1997). How we learn variation, optionality, and probability. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam* 21, 43-58.

Boersma, P. and C. Levelt (2000). Gradual constraint ranking learning algorithm predicts acquisition order. In the *Proceedings of the 30th Child Language Research Forum*, ed. E. Clark, p. 229-37. Stanford, CA: CSLI.

Clements, G. (1990). The role of the sonority cycle in core syllabification. In *Papers in Laboratory Phonology: Between the Grammar and Physics of Speech*, ed. J. Kingston & M. Beckman, 282-333. Cambridge: Cambridge University Press.

Levelt, C. and Van de Vijver (to appear). Syllable types in cross-linguistic developmental grammars. In In *Fixing Priorities: constraints in Phonological Acquisition*, ed. R. Kager, J. Pater & W. Zonneveld. Cambridge: Cambridge University Press.

Pater, J. and J. Barlow (to appear). Constraint ranking in child production and perception. *Journal of Child Language*

Phonological change of Taiwan Mandarin —Evidence from the Three Presidential Candidates' Speech

Chun-Mei Chen
University of Texas at Austin

This study focuses upon voice data from mass media to disclose the phonetic facts of modern Taiwan Mandarin and to identify its phonological change. The study consists of perceptual judgments and acoustic analyses of continuous speech. "Taiwan Mandarin" (TM) here refers to three typical varieties of Mandarin commonly spoken in Taiwan. Each candidate represents one unique kind of Taiwan Mandarin. Candidate Chen's speech represents the Mandarin variety which is heavily influenced by Taiwanese. Candidate Soong's speech represents the Mandarin variety of "Mainlanders", whose parents were born and raised in one of the provinces in Mainland China. In other words, candidate Soong's speech is influenced by at least one of the Chinese dialects spoken in Mandarin China. Candidate Lien's speech represents the variety that is employed by bilingual speakers of Mandarin and Taiwanese. Based on perceptual judgments, the phonetic variants of each candidate were identified. The acoustic analyses interpreted some but not all of the results obtained from perceptual judgments. Subtle tone variants were detected by perceptual judgments, rather than acoustic measurements. We built in the rules of phonological change to account for each variety of Taiwan Mandarin.

The results show that bilinguals and language learners as speakers are one of the sources of phonological change in Taiwan Mandarin. As far as the syllable-initial consonant is concerned, Chen's retroflex, fricatives, nasals, palato-alveolars were perceived deviant from the standard Mandarin Chinese. In the Medial glide category, Chen substituted [i] for [y] at the syllable medial position. Chen's insertion of Medial [u] is conditioned by the initial labial consonants. In the Rime category, Chen's pronunciations were perceived as either missing medial glide or replacing velar nasal with its alveolar counterpart. In the Tone category, Chen's second tone was perceived as the third tone. The contrast between Taiwanese and Mandarin and the influence of Chen's mother tongue can account for the phenomena mentioned above. Soong's speech was judged to be within the standard range, except for his lack of alveolar nasal consonant, which is the main difference between his dialect and Mandarin, and his variation of velar nasal ending. Lien's speech was judged to be standard, except for his velar nasal ending and tones. The rising feature of the second tone was lost in both candidate Chen's and candidate Lien's speech. Their second tone becomes the third tone in various phonetic environments. This could be attributed to the low starting point of the second tone in their speech. The third tone in the three varieties of Taiwan Mandarin has no rise. In other words, the phonological feature of the third tone in Taiwan Mandarin is only falling, rather than a complete falling-rising contour.

The study provides acoustic and perceptual evidence for the three varieties of Taiwan Mandarin. It also provides implication for language pedagogy. The second language learners of Mandarin Chinese have to realize the varieties of Taiwan Mandarin in order to enhance their listening comprehension of Taiwan Mandarin.

Perceptual reality of phonetically-driven phonology: Place assimilation and consonant cluster simplification at different prosodic boundaries

Taehong Cho & James McQueen
Max Planck Institute for Psycholinguistics

It has been proposed that some phonological patterns are best explained by the acoustic/perceptual characteristics of segments (see [1]). For example, alveolars tend to undergo place assimilation more often than labials, which are in turn more susceptible to place assimilation than velars. This asymmetry in place assimilation is arguably accounted for by the so-called Production Hypothesis: speakers make a greater effort to preserve perceptually stronger segments (velars) than perceptually weaker segments (alveolars), from which listeners benefit ([2,3]; cf. [4]). Such a phenomenon is also seen in tri-consonantal cluster simplification in Korean: for a C1C2C3 sequence, when C2 is optionally deleted, speakers delete /t/ most often and /k/ least often. This has been interpreted as implying that /k/ is kept more often because its perceptual salience is greater than /t/ ([5]). However, the perceptual reality of such interpretations have not yet been corroborated by speech perception data.

We evaluated the perceptually-grounded Production Hypothesis by investigating how the asymmetry in place assimilation and cluster simplification in Korean would influence phoneme recognition. 104 Seoul Koreans participated in a series of phoneme monitoring-restoration experiments in which they monitored for an *underlying* coda target in either an assimilated (/t/ or /p/ -> [k]/_k) or an unassimilated real word form (Exp.1); in either an illegally assimilated (/p/ or /k/->[t]/_t) or an unassimilated real word form (Exp.2). In another experiment (Exp.3), they had to monitor for C2 when it was present in a C2-kept word ({C1}C2C3, '{}'=a deleted phoneme) or restore C2 when it was absent in a C2-deleted word (C1{C2}C3). In order to examine whether and how the phoneme recognition process is constrained by prosodic structure, words with target phonemes were placed Accentual Phrase (AP) -initially and AP-medially (cf. [6,7])

Several findings have emerged. First, the speed and accuracy of recognition of the unassimilated targets were in general greatest for velars, intermediate for labials and smallest for alveolars in line with the hypothesized perceptual salience ranking: velars > labials > alveolars. More specifically, comparing the unassimilated targets with assimilated ones, the recognition of the unassimilated targets were faster and more accurate than assimilated ones only when the targets were labials (but not when the targets were alveolars) (Exp.1).

Second, comparing the illegally assimilated targets with the unassimilated ones, the recognition of the illegally assimilated targets (e.g., restoring /k/ when /k/-> [t]/_t) in real word forms was significantly slower and less accurate than that of the unassimilated targets. However, there was a significant interaction with the consonant type: the effect was significantly far greater for velars than for labials (Exp.2). This supports the hypothesis that velars are perceptually more salient than labials, such that the loss of (more salient) velars due to illegal assimilation is more harmful than that of labials, making the restoration of the former harder than the latter.

Third, for the cluster C1C2C3 which is simplified by deleting either C1 or C2, when the C2 targets in words were velars, the recognition of C2 was significantly faster for {C1}C2C3 items (with the velar target present) than for C1{C2}C3 items (with the velar target absent). Crucially, however, when the C2 targets in words were labials no such effect was found, suggesting that listeners benefit from the preservation of velars, but not from that of labials (Exp.3).

Finally, as for prosodic boundary effects, the differences in phoneme recognition/restoration due to segment type were significantly more robust AP-initially than AP-medially, indicating that listeners are more sensitive to phonological changes at a stronger prosodic boundary.

Overall, the results suggest that asymmetries in place assimilation and consonant cluster simplification in Korean are indeed perceptually grounded, supporting the view that phonetics can be a source of explanation for some phonological patterns. Moreover, we infer that the human speech recognition system is sensitive to the acoustic/perceptual properties of individual segments to a different degree, which is further modulated by the prosodic structure of a given language.

References

[1] Hume, E. & K. Johnson (2001) The Role of Speech Perception in Phonology. Academic Press.

[2] Jun, J. (1995) Perceptual and Articulatory factors in Place Assimilation: An Optimality Theoretic Approach. PhD dissertation, UCLA.

[3] Jun, J. (in press) Place assimilation. In B. Hayes, R. Kirchner and D. Steriade (eds.) The Phonetic Bases of Markedness. Cambridge University Press.

[4] Steriade, D. (2001) Directional asymmetries in assimilation: A directional account. In E. Hume & K. Johnson (eds.), The Role of Speech Perception in Phonology, pp. 219-250.

[5] Cho, T. (1999) Intra-dialectal variation in Korean consonant cluster simplification: A stochastic approach. CLS 35,43-58.

[6] Jun, S.-A. (1996) The Phonetics and Phonology of Korean Prosody: intonational phonology and prosodic structure, Garland Publishing Inc., New York : NY.

[7] Kim, S. (2003) The role of post-lexical tonal contours in word segmentation. In Proceedings of the 15th International Congress of Phonetic Sciences. Barcelona, Spain, 2003.

L2 re-categorization of an 'old' phonological contrast

Paola Escudero* & Paul Boersma†
Utrecht University and University of Amsterdam†*

This paper argues that the L2 development of similar sounds (i.e. sounds that are common to the native and the target language) is most accurately explained by a functional approach to L2 sound categorization, which hypothesizes a strong dependency of perception on the ambient production distributions.

Canadian English (CE) has a phonological contrast between the vowels /ae/ and /E/ (as in "bat" and "bet"). Canadian French (CF) has a similar /ae/-/E/ contrast (as in "bac" and "bec"). Acoustic analyses reveal that the French vowels differ from their English counterparts in F1 and duration. We recorded six speakers (3 male, 3 female) of each language producing each vowel in five CVC contexts, embedded in CE and CF carrier sentences. These 120 vowel tokens were analysed for F1 and duration. It turns out that CE uses both vowel duration and F1 to signal the phonological contrast, while the CF contrast is signalled by F1 differences only. Tokens of CE /ae/ turn out to have higher F1 values and longer durations than tokens of CF /ae/. We used these productions to measure CE and CF monolingual perception, CE-to-CF cross-language perception, and L2 perception of CE learners of CF.

Eight monolingual CE and eight monolingual CF listeners had to classify all 120 tokens into their native /ae/ and /E/ categories, while being told (incorrectly) that all the stimuli were tokens of their native language. When categorizing their 60 native vowels, both groups of listeners were above 90% accuracy and showed a category boundary that matched the production distributions of their specific native language. We conclude that listeners follow a maximum-likelihood strategy, i.e., they tend to classify every token as the vowel phoneme that was most likely to have been intended by the speaker.

We then examined the cross-language perception of the 60 CF vowel tokens by the eight monolingual CE listeners. They turned out to classify 36% of the CF /ae/ tokens as /E/ and showed a category boundary that resembled their native language perception (and, therefore, their own production distributions). From this we conclude that monolingual cross-language perception is determined by the perception of the native language.

Our L2 study, finally, investigated 23 CE learners of CF. Their L2 proficiency level (8 beginners, 8 intermediates, 7 advanced) was independently established on the basis of a language background questionnaire. These learners had to classify the 60 CF vowel tokens as CF /ae/ or /E/. The average beginning learner categorized CF /ae/ correctly 71% of the time. Since this number is close to the 64% for English monolingual cross-language perception seen above (64% of CF /ae/ is classified as /ae/), the beginners' behaviour can be explained by hypothesizing that they re-use their native CE /ae/ category as their L2 CF /ae/ category, and use their native monolingual /ae/-perception strategies to classify the CF tokens. The intermediate group showed improvement (84% correct), and the advanced learners were very close to the native French listeners (91%). Importantly, the learners' L1 perception did not change in the process.

Taken together, these perceptual and acoustic results suggest that a maximum-likelihood approach is the most accurate in explaining the native, cross-language and L2 categorization patterns. We will show that a specific phonological formalization of this functional approach (namely, with language-specific Optimality-Theoretic perception grammars) can successfully handle the phenomena attested in our experiments: language-specific perceptual mappings that

resemble the ambient distributions, re-use of L1 categories and perceptual mappings in cross-language categorization, and the subsequent optimisation of separate perception systems for L1 and L2 (i.e. developmental changes in L2 did not affect the L1 system).

Does Prosodic Phrasing Correlate with Wh-scope in Japanese?

Masako Hirotani
University of Massachusetts

In Japanese wh-questions with scopally ambiguous wh-phrases, Hirotani [1] found that listeners have a strong preference for assigning the embedded scope to the wh-phrase when a major phrase boundary appears after the embedded Q-marker (Emb-Q). This preference disappeared in the absence of such boundary. This finding contrasted with the previous proposals about the syntax-phonology interface [2][3]. The current study investigates whether speakers produce distinct prosodic phrasings for different scopal interpretations of wh-phrases.

The unambiguous Japanese wh-questions in (1) were tested in the present production experiment: With the matrix Q-marker *nokai*, the wh-phrase takes the embedded scope, as in (1a) while with the Q-marker *ndai*, the matrix scope is assigned to the wh-phrase, as in (1b). These two forms were produced by 4 native Tokyo speakers. Short and long forms were tested. The short forms are given in (1). In the long forms, either temporal or locative prepositional phrase(s) were added following both the wh-phrase and the Emb-Q.

(1) a. John-wa Mary-ga **nani**-o katta-ka kiita-**nokai**?
John-TOP Mary-NOM what-ACC bought-Q asked-Q
'Did John ask what Mary bought?' (Embedded Question)

b. John-wa Mary-ga **nani**-o katta-ka kiita-**ndai**
John-TOP Mary-NOM what-ACC bought-Q asked-Q
'What did John ask whether Mary bought?' (Matrix Question)

The following measurements were taken to investigate whether pitch resetting occurred after the Emb-Q: (i) The peak F0 value subtracting that of the L-tone on the Emb-Q from that on the H-tone of the word immediately following the Emb-Q (Post-Q-Word) (I call this X), (ii) the duration of the Emb-Q (Y), and (iii) the length of pause between the Emb-Q and the Post-Q-Word (Z).

In the long forms, pitch was always reset after the Emb-Q for both types of questions: (i) X was significantly larger than 0 for all speakers and for 3 speakers X did not differ significantly between the two types of questions [Mean X of all speakers: 143 Hz (Embedded Question) vs. 142 Hz (Matrix Question)], (ii) Y was long compared to that of other words in the sentence and did not differ significantly between the two types of questions [Mean Y of all speakers: 215 ms (Embedded Question) vs. 150 ms (Matrix Question)], and (iii) A pause was likely to be inserted and for two speakers, Z was significantly longer for the embedded questions than the matrix questions [Mean Z of all speakers: 266 ms (Embedded Question) vs. 200 ms (Matrix Question)]. The presence of pitch resetting after the Emb-Q in the long forms is likely to be due to the left edge of the maximal projection of the inserted adjunct phrase following the Emb-Q [4].

The results for the short forms were much more complicated. In the embedded wh-questions, pitch resetting generally occurred after the Emb-Q-marker: For 3 speakers, X was significantly different from 0 [Mean X of all speakers: 52 Hz (Embedded Question)]. In the matrix wh-questions, the occurrence of pitch resetting was variable: X was not significantly larger than 0 for 3 speakers [Mean X of all speakers: 44 Hz (Matrix Question)]. The results of the analysis with other measurements were consistent with this pattern of results: For the majority of the speakers, Y did not differ significantly between the two types of questions, and

no significant difference was found in Z. This variability in the matrix wh-questions may be due to the variability in the focus structure required for the sentence: The embedded wh-questions require either presentational or contrastive focus after the Emb-Q while the matrix questions do not. Focus alignment with a MaP can explain the pitch resetting in the embedded wh-questions [5][6].

The current study in combination with Hirotani [1] suggests that the prosodic phrasing listeners prefer does not always match with the prosodic phrase speakers produce.

References

[1] Hirotani, M. 2003. "Prosodic Effects on the Interpretation of Japanese Wh-questions". In L. Alonso-Ovalle (Ed.), University of Massachusetts Occasional Papers in Linguistics 27: On Semantic Processing. University of Massachusetts, Amherst: GLSA, 117 - 137.

[2] Deguchi, M. and Y. Kitagawa. 2002 "Prosody and Wh-questions". In M. Hirotani (Ed.), Proceedings of North East Linguistic Society, Volume 32, University of Massachusetts, Amherst: GLSA, 73 - 92.

[3] Ishihara, S. 2002. "Invisible but Audible Wh-scope Marking: Wh-constructions and Deaccenting in Japanese". In L. Mikkelsen and C. Potts (Eds.), Proceedings of West Coast Conference on Formal Linguistics, Volume 21. Somerville, MA: Cascadilla Press, 180 - 193.

[4] Selkirk, E. and K. Tateishi. 1991. "Syntax and Phonological Phrasing in Japanese". In C. Georgopoulos and R. Ishihara (Eds.), Interdisciplinary Approaches to Language, Essays in Honor of S.-Y. Kuroda. Dortrecht: Kluwer.

[5] Nagahara, H. 1994. Phonological Phrasing in Japanese, Ph.D. dissertation, University of California, Los Angeles.

[6] Sugahara, Mariko. 2003. The Phonetics and Phonology of Downtrends in Tokyo Japanese Intonation, Ph.D. dissertation, University of Massachusetts, Amherst.

Kitagawa, Y. and J. D. Fodor. 2002. "Default Prosody Explains Neglected Syntactic Analyses in Japanese". In W. McClure (Ed.), Japanese/Korean Linguistics, Volume 12, Stanford University: CSLI Publication.

Maekawa, K. 1997. "Nihongo Gimosi Gimonbun-no Intoneesyon [Intonation of Japanese Wh-questions]". In Spoken Language Working Group (Ed.), Speech and Grammar. Tokyo: Kuroshio Publisher, 45 - 54.

Effects of L1 phonology and orthography on L2 perception: The use of Korean Hangul in perception of English voicing contrasts

Byung-jin Lim
University at Albany, State University of New York

There are two kinds of asymmetries sensitive to syllable position regarding phonology and orthography in Korean. Firstly, only syllable initial position preserves phonemic contrasts (a three-way distinction in Korean stops: '? ' lax stop vs. '? ' tense stop vs. '? 'aspirated stop) while the phonemic contrasts neutralize into lax stops in syllable final position (i.e., syllable final neutralization). Secondly, the syllable final neutralization is only captured in phonology, not in orthography ("leaf" (?) /ip^h/ ✍ [ip]; "mouth" (?) /ip/ ✍ [ip]). However, English voicing contrasts are not position-sensitive, and they are well represented in both phonology and orthography in English (**b**at vs. **p**at; mo**b** vs. mo**p**). Thus, a possible problem may arise when Koreans perceive English voicing contrasts especially in syllable final position and represent them with their own orthography, *Hangul.*

Orthographies differ in the degree of complexity in the relationship between spelling and sound: 1) shallow and 2) deep orthographies. In shallow orthographies, the correspondence between print and sound is transparent and regular (e.g, Serbo-Croatian and Italian). On the other hand, deep orthographies characterize languages in which the mapping between print and sound is opaque in that there are many inconsistencies in spelling-to-sound rules (e.g., Hebrew and Chinese). Korean is chosen since it is relatively transparent in terms of mapping sounds onto letters such that the relationship between spoken and written representation of speech is fairly straightforward. Given that Korean has no phonemic voicing distinction, Korean will be a good testing ground for investigating how English voicing contrasts are perceived, and thus represented in its writing system.

This study uses perceptual tasks in repetitive speech (Stetson, 1951) where repeated singleton coda consonants tend to be perceived as onset consonants as the rate of repetition increases. The current perception study employs a similar perceptual task (de Jong 2001a, 2001b) using four types of syllables (/bi/, /pi/, /ib/, and /ip/) as stimuli recorded by two native speakers of American English at various rates controlled by a metronome. There were 336 stimuli for the perception test, and two groups of 20 Koreans, who were recruited in Seoul, were asked to identify voicing contrasts of stops. There were four possible choices labeled with 1) English orthography ('bee', 'pea', 'eeb', 'eep') or 2) Korean orthography ('? ' /pi/, '? ' /p^{h}i/, '? '/ip/, '? ' /ip^h/) from which listeners would make their best choice for each stimulus by clicking a mouse on a computer screen. Eighteen Americans also participated as controls.

Identifications of voicing contrasts from stimuli /pi/ and /bi/ indicate that there are tendencies Korean /? / substituting for English /p/ and /? / for /b/ respectively. From the stimulus /ip/ Korean listeners who were tested with English orthography show a similar identification pattern to native English controls. On the other hand, Korean listeners who were tested with Korean orthography show about a chance level of identification at slow speech rates and they shift toward '? ' as speech rate increases. It seems that the Korean display causes more confusion to Korean listeners in identification of voicing than does the English display. In addition, the Korean orthography reflects frequency of characters contributing to bias toward '? ' (? /b/) over '? '(? /p/) as perceptual responses of voicing contrasts. It is also suggested that

the fact that '? ' tends to substitute for English /b/ orthographically biases Korean listeners toward [b] of post-vocalic English stops.

References

de Jong, K. J. (2001a). Effects of syllable affiliation and consonant voicing on temporal adjustment in a repetitive speech production task. *Journal of Speech, Language, and Hearing Research*, 44, 826-840.

de Jong, K. J. (2001b). Rate-induced resyllabification revisited. *Language and Speech*, 44 (2), 197-216.

Stetson, R. H. (1951). *Motor Phonetics*. Amsterdam: North-Holland.

The Phonetics and Phonology of Front Vowels in Staten Island English: when the traditional descriptions and the facts do not agree

Bruce Morén
University of Tromsø

According to the 2000 U.S. Census, there are over 8 million people living within the five boroughs of New York City, and the borough of Richmond (a.k.a. Staten Island) has just over 440,000 inhabitants. Until the late 1960s, Staten Island was fairly isolated, with travel to and from the rest of the city possible only by ferry or by crossing a bridge to New Jersey and then a bridge or tunnel to Manhattan. Because of this relative isolation, Staten Island had its own, distinct, dialect. On November 21, 1964, the island's cultural and linguistic identity changed forever as the Verrazano-Narrows Bridge opened a direct route to Brooklyn and the rest of the city. Since that time, the sociolinguistics of the island have become increasingly complex with a number of dialects merging and emerging as people have moved from more urban metropolitan areas (particularly Brooklyn) to the more rural (now suburban/urban) Staten Island. As a result, a unique dialect is slowly being lost as the population becomes more homogenized with the rest of the city.

In this talk, I present the results of a pilot study investigating the front vowel inventory of native speakers of the dialect spoken on the south shore of Staten Island. This dialect is interesting from both a sociolinguistic and phonetics-phonology perspective since it suggests that phonologists working on dialectal variants should be wary of relying too heavily on traditional language descriptions that have not been supported by instrumental data. In addition, it suggests that care is needed not to over-generalize descriptions or analyses across sub-dialects, even for robustly reported sociolinguistic markers. In fact, Staten Island English clearly shows that different dialects can have very similar overall phenomena, and yet differ significantly in the details pertaining to those phenomena.

This talk begins by reviewing the three predominant, and conflicting, phonological analyses regarding the distribution and nature of the two low front vowels found in Metropolitan New York English - a phenomenon usually described as 'æ-tensing'. The first analysis describes both low front vowels as 'short a' and their relationship as purely allophonic (Cohen 1970). The second analysis describes the two vowels as being in a semi-allophonic relationship, where /æ/ is short and lax, while /E/ is long, tense and diphthongized (Labov 1981). In the third analysis, the phonological patterning of the vowels is used to argue that they have distinctive length, which is subject to context-specific neutralization to long. In direct contradiction with the second analysis, it claims that /æ/ is long and /E/ is short (Morén 1999).

The second part of the talk presents the results of an acoustic study undertaken to see which of the conflicting phonological analyses is closest to the phonetic facts. In this study, the relative durations of all the Staten Island front vowels, as well as their relative formant structures, were measured in a variety of phonological contexts. Somewhat surprisingly, the results show that none of the analyses found in the literature provides an accurate or adequate picture of æ-tensing for this dialect. Rather, we are left with the conclusion that the low front vowel that is traditionally described as 'tense' is actually lax, the 'lax' vowel is actually tense, and despite the fact that at least one of the low front vowels is usually described as short, both vowels are actually phonetically and phonologically long.

Are functional constraints active synchronically?

Geoffrey Stewart Morrison
University of Alberta

This paper reports on an experiment to test whether a functional phonological constraint is active synchronically.

Functional approaches to phonology such as Flemming (1995), Boersma (1998), Kirchner (1998), and Steriade (1999) claim that constraints penalising articulatory effort and perceptual confusion are major driving forces in phonology. There is, however, a debate as to whether the patterns proportedly accounted for by functional constraints are due to active synchronic constraints or whether they are a vestige of diachronic change. At one extreme Hale and Reiss (2000) argue that patterns observed in the output of the phonological system are due to the existence of these patterns in the input, and that there is no need to posit functional constraints on phonology to account for said patterns. They argue that although articulatory and perceptual factors may influence diachronic change, they are irrelevant to phonology and should therefore be excluded from phonological theories. In contrast, Kirchner (2001) argues that since functional phonetic constraints are necessary to account for allophonic variation, and the same constraints can also account for phonemic variation, there is no need to posit a distinct system to account for the latter.

Hale and Reiss (2000) claim that a neutralisation may emerge if children learning their first language fail to perceive a contrast between two sounds in a context in which they are acoustically similar. Since this process is a phonological change and the cause of the change is posited to be perceptual confusion, I would argue that it should be formalised in a functional phonology framework. The change would be diachronic but the functional constraint would also be active synchronically, at least in the stage during which the children learn the phonology. I hypothesise that if functional constraints are active synchronically, then phonetically-natural phonemic alternations will be easier to learn than phonetically-unnatural phonemic alternations. If, however, functional constraints are not active synchronically, then phonetically-natural and phonetically-unnatural phonemic alternations will be equally easy or difficult to learn.

The functional constraint tested in the present study is an articulatory constraint resulting in an alternation in which fricatives appear between two low vowels, and plosives appear between two vowels when at least one of the vowels is non-low. To make a plosive closure between two low vowels the jaw or tongue must be displaced and returned a greater distance than between two non-low vowels or one low and one non-low vowel. The greater displacement requires a greater articulatory effort. The constraint would be an instantiation of Kirchner's (1998) LAZY or Boersma's (1998) *EFFORT, e.g., *GESTURE(stop closure / a_a). The hypothesis is tested using a laboratory-learning paradigm: Two balanced groups of participants are trained and tested via aural and picture presentations of morphological alternations. One group is trained on a phonetically-natural alternation in which the first segment of the base is a fricative when it is between two low vowels (i.e, when the last segment of the prefix and the second segment of the base are both low vowels) but is a plosive otherwise. The other group is trained on an identical set of bases and prefixes but with a phonetically-unnatural alternation, the first segment of the base is a plosive when it is between two low vowels but is a fricative otherwise. If the group trained on the phonetically-natural alternation learns the alternation faster than the group trained on the phonetically-unnatural alternation, then this will be taken as

evidence in support of the hypothesis that the functional constraint is synchronically active. Results of the experiment will be presented at the conference.

References

Boersma, P. (1998). Functional phonology: Formalising the interaction between articulatory and perceptual drives. Doctoral dissertation, University of Amsterdam. The Hague: Holland Academic Graphics.

Flemming, E. (1995). Auditory features in Phonology. Unpublished doctoral dissertation, University of California Los Angeles.

Hale, M., and Reiss C. (2000). Substance abuse and dysfunctionalism: Current trends in phonology. Linguistic Inquiry, 31, 157-169.

Kirchner, R. (1998) An effort-based approach to consonant lenition. Unpublished doctoral dissertation, University of California Los Angeles.

Kirchner, R. (2001). [Review of A. McMahon (2000) Change, chance, and optimality]. Phonology 18, 427-433.

Steriade, D. (1999). Phoneics in phonology: The case of laryngeal neutralization. In M. K. Gordon (Ed.), Papers in Phonology 3, UCLA Working Papers in Linguistics 2, 25-146.

Convergence or Divergence? Accommodation in the Vocalic System of Texan English

Ana Sánchez Muñoz
University of Southern California

This paper examines the type of speech modification or "**accommodation**" (Trudgill 1986) that affects the vowel production of speakers in the context of contact between two varieties of the same language. In this case, the vowel production of two Texan speakers of English is examined in order to test for any significant qualitative difference in their vowel production based on such extralinguistic factors as social identity.

Sociolinguistic research has demonstrated that there is a strong correlation of independent social variables with phonetic variants. At the same time, it has long been observed that upon contact with a new variety, some speakers adopt aspects of it (**convergence**), while others do not (**nonconvergence** or **divergence**). It seems that the phenomenon by which speakers attempt to conform to the speech behavior of those around them is highly dependent on social aspects (Giles, Taylor, and Boorhis 1973). **Language attitude** is one social factor that plays an important role in linguistic convergence. Dialect diversity is a linguistic universal, and the fact that some dialects are more highly praised than others is socially universal. Therefore, speakers can make conscious efforts to alter their own speech in the direction to what they consider the 'standard' or more prestige variety of their language. This happens at the expense of their own less prestigious regional variety (Baugh 1993: 177).

The **hypothesis** of this study is related to language attitude as a main driving force for language variation in the speech of some individuals (Labov 1972; 1994). In this paper, I claim that phonetic variation may occur among speakers who move from a non-standard linguistic region to a more standard-like area. In particular, changes in these speakers' vowel production may be observed by acoustic analyses across different linguistic situations. By "standard", I specifically refer to the subject's own perception of what is a more prestigious variety of a language, rather than to a general accepted view of the term as the normative language. This hypothesis is based on the principle of the **adaptive function of linguistic diversity** (Labov 1972: 323-45). In order to investigate this type of standard-nonstandard phonetic variation, this paper examines the speech of two Texan speakers of English and provides an acoustic analysis of their vowel production. The following features are explored: (a) the monophthongization of /ai/ and /ei/; (b) the lowering of /iʔ/; (c) the merger of /e/ and /i/ before nasals; (d) pre-/l/ mergers.

The **results** obtained from three experiments show that there is indeed within speaker variation in vowel production depending on the linguistic situation in which the subjects are involved. These results suggest that phonetic variation within speakers can be driven by social factors such as the association of a linguistic form with membership in that social group (local adaptivity) or the reverse (global counteradaptivity) (Chambers 1985; Tabouret-Keller 1985).

References

Baugh, J. 1993. Adapting Dialectology. In Dennis R. Preston (ed.), *American Dialect Research*. Amsterdam/Philadelphia: John Benjamins.

Chambers, J.K. 1985. Social Adaptiveness in Human and Songbird Dialects. *The Behavioral and Brain Sciences* 8:85-100.

Giles, H., Taylor, D., and Boorhis, R. 1973. Towards a theory of social accommodation through language: Some Canadian data. *Language in Society* 2: 177-192.

Labov, W. 1972. *Sociolinguistic Patterns*. Conduct and Communication 4. Philadelphia: University of Pennsylvania Press.

------1994. *Principles of Linguistic Change*. Oxford : Blackwell.

Labov, W., Malcah Yaeger, and Richard Steiner. 1972. *A Quantitative Study of Sound Change in Progress*. Philadelphia. U.S. Regional Survey.

Tabouret-Keller, A. 1985. *Vernacular literacy: a re-evaluation*. Oxford : Clarendon Press ; New York : Oxford University Press.

Trudgill, P. 1986. *Dialects in Contact*. Oxford: Blackwell.

Exploring performance-based predictors of phonological judgments in Mandarin

James Myers & Jane Tsay
National Chung Cheng University

Since judgments measures of competence are a form of performance, an understanding is needed of the relation between judgments and normal language processing (see also Schutze, 1996). To begin to address this in phonology, off-line phonological acceptability judgments were taken from native speakers of Taiwan Mandarin, who also provided more traditional performance measures involving perception, production, and memory, on the same materials. Analysis consisted of modeling the judgment scores with scores from the other three tasks as predictors.

Mandarin allows more of the "logically possible morpheme space" to be tested than in English, due to the small syllabary and predominance of monosyllabic morphemes. The portion we tested consisted of all 384 combinations of 8 onsets, 4 nuclei, 3 codas (including no coda), and 4 tones. Of these, 222 are lexical items, 80 appear to be systematic gaps, and 82 appear to be accidental gaps. This set was divided into two subsets with equal numbers of lexical items, systematic and accidental gaps, and tokens of segmental syllable types. Participants first performed a judgment task, in which syllables from their assigned subset were presented auditorily and also visually in a phonetic orthography familiar in Taiwan; unlimited time was given to judge degree of Mandarin-likeness on a six-point scale. This task was then followed by production and perception tasks (order counterbalanced). In the production task, participants were shown their assigned subset of syllables in phonetic orthography and had to speak them as quickly as possible into a microphone voice key. In the perception task, participants heard their assigned syllables in noise and had to type the phonetic forms as accurately as possible. The final task was a memory task, in which all participants were presented auditorily and visually with all syllables and had to identify the previously presented ones quickly and accurately.

The overall pattern of judgments was consistent with what has been found in previous studies (Wang, 1998; Myers, 2002). The mean score for real words (5.4) was much higher than that for nonwords, but that for accidental gaps (2.9) was also significantly higher than for systematic gaps (2.4), showing that phonology played a role in addition to lexicality. This conclusion is also supported by the fact that the mean scores for real-word segmental syllable types were well correlated across the two subsets ($r(63) = 0.44$, $p = 0.0003$), showing consistent differences in judgments even among lexical syllables. A stepwise regression model of by-item mean judgment scores then revealed that the most important predictors were accuracy rates for the production task (coefficient = 2.59) and accept responses in the memory task (3.02), with small but significant roles for latencies in these two tasks as well (respectively -0.009, -0.006). Performance in the perceptual task and reject responses in the memory task played no role in the model. A similar model emerged when only lexical syllables were analyzed, but a striking difference was seen with nonwords, for which the chosen model gave a significant role to perception and reject responses (both latency and accuracy) in the memory task, but not to accept responses or production accuracy. For all models, however, the relatively large intercept (e.g., 10.31 for the all-syllable model) and relatively low R-squared (0.563) imply that none succeeded in capturing all of the variability in the judgment scores.

These findings lead to two conclusions about competence measures. First, though judgments are affected by aspects of normal processing, the incompleteness of the regression

models and the greater importance of accuracy scores over reaction times as predictors imply that judgments of linguistic knowledge involve much more than the on-line use of this knowledge. Second, the "judgment function," whatever it turns out to be, is not uniform: different processing strategies are used depending on the lexicality of the target item.

References

Myers, J. (2002). An analogical approach to the Mandarin syllabary. Journal of Chinese Phonology, 11, 163-190.

Schutze, C. T. (1996). The empirical base of linguistics: Grammaticality judgments and linguistic methodology. Chicago: University of Chicago Press.

Wang, S. H. (1998). An experimental study on the phonotactic constraints of Mandarin Chinese. In B. K. T'sou (Ed.), Studia Linguistica Serica (pp. 259-268). Language Information Sciences Research Center, City University of Hong Kong.

Polish sibilants: studying perceptual cues to explain the directionality of sound change

Pawel M. Nowak
University of California-Berkeley

Polish is one of the few languages that have sibilant fricatives and affricates at three contrasting places of articulation: dental, retroflex and alveopalatal. There is abundant anecdotal as well as experimental evidence that speakers of other languages find the retroflexes and the alveopalatals very similar to each other and difficult to distinguish (Lisker 2001). The similarity of these two sound types is also clear from spectral evidence. If sound change originates in misperception by the listener, as claimed, among others, by Ohala (1993) and Blevins (to appear), one would expect that when a system like this is simplified to one with only a twofold contrast, those two sound types should merge. However, the numerous Polish dialects that now lack the retroflexes have merged them with the acoustically more different dentals instead. Interestingly, parallel dialectal phenomena are observed in Mandarin, an unrelated language with a very similar inventory of sibilants (Duanmu 2000).

In this paper, we claim that these facts are only an apparent counterexample to perception-based models of sound change, because perceptual cues provided by formant transitions of the vowels adjacent to these sibilants are more important in recognizing their POA than the frication noise itself. Specifically, alveopalatals crucially require strongly lowered F1 and strongly raised F2 for adequate identification. Dentals and retroflexes, on the other hand, are similar in having virtually flat formant transitions. In this light, the directionality of the attested merger is no longer surprising, since it can be said to involve two sound types that are most similar in what is perceptually the crucial respect.

The reminder of this abstract describes the experimental design used in the study. A female native speaker of Polish recorded nonce-words of the form eCe, aCa and uCu, where C stands for any of the three contrastive voiceless sibilant fricatives. Two tokens of each were selected and divided into the consonant part and the vocalic parts. Next, these were combined and modified in seven different ways:

1. same-splicing (the consonant occurring in the same vocalic context as originally)
2. reversed retroflection (dentals and palatals occurring between vowels that originally surrounded a retroflex consonant; retroflex fricative occurring in the dental vowel context)
3. reversed palatality (retroflexes and dentals occurring in the original palatal vocalic context, the palatal fricative in the dental context)
4. reversed palatality – first vowel only
5. reversed palatality – second vowel only
6. vowel transitions removed
7. vowel edges removed

As a result, 63 words were created (7 conditions x 3 different vowels x 3 different fricatives). Six tokens of each were used to create a randomized set of 378 stimuli, which were subsequently presented to 7 adult native speakers of Polish in a forced choice classification task. The recognition of alveopalatals was dramatically reduced in conditions 2, 3, 5 and 6, to below-chance levels, showing that formant transitions are essential for the perception of this sound type. (Interestingly, the absence of transitions in just the preceding vowel (condition 4) did not seem to have a strong effect on the identification rate, indicating that the second vowel plays a much more important role.) Retroflexes and dentals were less strongly affected by cross-splicing and modification. In statistical terms, a two way ANOVA (with splicing condition and place of

articulation as factors) revealed a highly significant effect for both factors, as well as a highly significant interaction.

References

Blevins, Juliette (to appear) Evolutionary Phonology. Cambridge: Cambridge University Press.

Duanmu, San (2000) The Phonology of Standard Chinese, Oxford: Oxford University Press.

Lisker, Leigh (2001) "Hearing the Polish sibilants [s š s]: Phonetic and auditory judgments" In Grønnum, Nina and Jørgen Rischel (eds.) Travaux du Cercle Linguistique de Copenhague XXXI. To honour Eli Fischer-Jørgensen. Pp. 226-238. Copenhagen: C.A. Reitzel.

Ohala, John (1993) "The phonetics of sound change". Charles Jones (ed.) Historical Linguistics: Problems and Perspectives. London and New York: Longman.

The importance of social factors in the loss or recovery of phonological units

Miren L. Oñederra
UPV-EHU, University of the Basque Country

The standardization of the Basque language can be said to be on its way. Its main grammatical foundations were established by the Academy in 1968. The need or the justification for pronunciation norms was not a clear matter, but finally the demands of mass-media professionals were satisfied. Norms for the pronunciation of formal Standard Basque were published in 1998. The variable application of those norms in the news broadcasting of Basque television and radio provides us with a rich corpus of data on several phonological issues, such as perceptual awareness and social preferences.

In order to study to which point choices in pronunciation are conscious, the opinion of TV and radio speakers and their language tutors were taken into account. Quite clearly the choice of basic phonological units and their concrete phonetic realization constitute a pattern of identification for the speaker, and we found that the resistance to follow a given norm or the enthusiasm to obey it are based on conscious "ideological" choices. Some of them significantly coincided with the views expressed by linguists and academicians in the discussions leading to the 1998 norms.

Factors such as the identification of a given sound with the influence of Spanish play a crucial role in the rejection of that sound. The case of the voiceless fricative velar [x] is a good example. That sound is successfully being substituted in the pronunciation of many reporters by the voiced palatal stop or fricative (contextual variants). Other dialectal realizations are also accepted, as far as they do not resemble [x]. Paradoxically [x] is completely regular as the native version of e.g. English [h] in proper names and loanwords, or in other native forms. This proves that social factors may be of much import in the shaping of a given phonological system, but cannot eliminate an already existing basic unit.

Social factors of prestige also explain the pejorative attitude of Basque speakers towards the ongoing neutralization of the opposition between lateral palatal and central palatal consonants, which is said to be of Spanish origin. Many speakers who do not pronounce the lateral ignore that it is so. Once they are made conscious, desperately try to learn it. The fact that those speakers do not perceive the difference between the lateral and the nonlateral palatals makes the enterprise hopeless.

However we cannot say that, once lost, a basic unit will never make it back into the system. That hypothesis seems to be invalidated by the apparent ease with which a considerable number of speakers get back the opposition between apical and laminal voiceless sibilants. The opposition is an exclusive feature of Basque vs. Spanish but was long ago lost in large Basque areas. Being more audible than the lateral-non lateral opposition, it is a better target for recovery, in spite of the fact that its loss is socially better accepted (because of its longer tradition and/or because the ghost of Spanish influence has never been clearly formulated).

We therefore conclude that, given the social motivation for it, lost sound units can be "recovered" by the system, only if perceptual prominence is enough. However stigmatized units can be reduced but not completely eliminated.

Articulatory weakening of an apical obstruent: lambdacized and rhotacized forms in Italian dialects

Michela Russo
Univ. de Paris 8/UMR 7023-CNRS

In some Neapolitan dialects we find the phenomena of rhotacism and lambdacism. That is, we find the occurrence of?[l] and [r] where a phonemic /d/ is assumed in the lexical representation. In the case of rhotacism, this is a special form of a general phenomenon which is characterized by the replacement of an apical obstruent by /r/.

Rhotacism is a well-known diachronic process, but it is also observed synchronically, for example in non-British variants of English such as American and Australian (known in these languages as /t/ and /d/ "flapping"), and in more local dialectal variants of British English (Liverpool and other Northwestern accents). Lambdacism appears to be less common, but it is also documented in Moselle-Franconian German.

The aim of this paper is firstly to describe the phonotactic and contextual conditions under which lambdacism and rhotacism have been observed within a recorded corpus of spontaneous speech, and secondly to suggest some phonetic explanation and phonological generalisation for the patterning we have found.

Rhotacism is plausibly considered to be the result of an articulatory weakening or shortening of an apical obstruent. From an articulatory phonetic standpoint, this is a convincing interpretation, since the shortening of the apical closure or stricture results in a very brief contact between tongue tip and the dental-alveolar region. Perceptually too, the interpretation is convincing: Romero (2001) has shown experimentally that, across languages, obstruent shortening results in an [r] percept, a necessary precondition for the perceptual re-interpretation and subsequent re-encoding of the weakened obstruent as /r/ that underlies diachronic rhotacism (cf. Barry & Russo 2002; 2003).

The case of lambdacism is not only less common, it is also less straightforward phonetically. The feature of laterality, which is the chief defining articulatory property of [l] vs. [d] (lambdacism appears to be restricted to /d/ replacement), cannot be seen *per se* as a natural gestural result of weakening the articulatory effort in apical stop production. However, [l] and [r] do share a property of their articulatory configuration which may help to explain the alternation between lambdacism and rhotacism in related dialects such as those under scrutiny in this study. The shared property can perhaps best be described as a "lack of tongue-blade raising", or at least a lack of contact between tongue blade and the front upper molars (cf. Barry 1997).

It would seem plausible that part of the "weakening" that underlies both rhotacism and lambdacism is a lack of the tongue-blade raising that is required to make side contact for apical obstruents. What might differentiate the tapped [r] from the lateral percept is the relatively shorter duration of the gesture, or in other terms, the ballistic tongue-tip movement vs. a more controlled raising (cf. Barry & Russo 2002; 2003).

Thus we do not claim that the reductions leading to [l] and [r] are *articulatorily* identical. The perceptual re-interpretation and consequent re-encoding of the weak /d/ as [l] or as [r], which is the prerequisite *sine qua non* of lambdacism and rhotacism probably has a substantive basis - gesturally and acoustic. We do, however, argue that the reduction of time and articulatory precision - probably without the tongue-blade raising required for a proper coronal closure leads to a percept that can be re-interpreted phonemically as [l] or [r].

Acoustic verification was based on the short (ca. 20ms) wideband attenuation of the acoustic signal in the case of [r] and on the local attenuation in the spectrum around 2 kHz indicating the zero or anti-resonance characteristic of a lateral. Secondary evidence in the case of [r] is a lowered F3 which sometimes accompanies the production of [r], particularly in front-vowel environment, and in the case of [l], the wideband impulses that reflects the "tissue noise" that sometimes occurs at the formation and closure of the lateral passages.

References

Barry, W. (1997): Another [r]-tickle. Journal of the International Phonetics Association, Vol.27.1-2, pp.35-45.

Barry, W. & Russo, M. (2002): Lambdacismo e rotacismo in area napoletana: realizzazione fonetica e posizione fonologica. R.Pompoli / E. Carletti / P. Fausti / F. Pedrielli / A. Peretti / F. Pompoli, N. Prodi (eds.), 29º Convegno Nazionle dell'Associazione italiana di acustica, Ferrara, 12-14 giugno 2002, Padova: Arti grafiche padovane, pp. 377-382.

Barry, W. & Russo, M. (2003): Lambdacismo e rotacismo nelle varietà di Ischia e Pozzuoli (NA): analisi acustico-uditiva delle varianti lenizzate dell'occlusiva sonora. P. Cosi / E. Magno Caldognetto / A. Zamboni (eds.), Voce Canto Parlato. Studi in onore di Franco Ferrero. Padova: Unipress, pp. 37-42.

Catford J.C. (2001): On Rs, rhotacism and paleophony. Journal of the International Phonetics Association, Vol.31.2, pp.171-185.

Coleman, J. (1998): Cognitive reality and the phonological lexicon: a review. Journal of Neurolinguistics, Vol. 11, pp. 295-320.

Harrington, J. & Cassidy S. (1999): Techniques in Speech Acoustics. Dordrecht: Kluwer Academic Publishers.

Romero J., (ms.) Temporal reduction effects in diachronic change; Rhotacism. Conference on the Phonetics-Phonology Interface, ZAS, Berlin, 11-13 October 2001.

The influence of Gujarati and Tamil L1s on Indian English

Caroline Wiltshire & James D. Harnsberger
University of Florida

English as spoken as an L2 in India has developed distinct sound patterns in terms of both segmental and prosodic characteristics. Indian English (IE) is primarily acquired in a classroom setting, beginning in the first grade and continuing through university; thus, even fluent IE speakers acquire this language in addition to an L1. We investigate the differences between two L1 groups to evaluate to what extent IE accents are based on a single target phonological-phonetic system ("the IE model" being taught) and/or influenced by transfer from L1. We examine phonetic and phonological variation in the L2 English of L1 Gujarati and Tamil speakers, and find L1 transfer effects in both segmental and suprasegmental properties.

Data was collected from 10 proficient IE speakers in Hyderabad, India: 5 Gujarati L1 and 5 Tamil L1, ages 18-24, with 4 females and 1 male for each L1. Each speaker was recorded reading a word list, isolated sentences, a short passage, and a series of short dialogues, and then answering questions in a language background interview. Preliminary results are presented for vowels, consonants, and intonation.

A set of keywords containing the vowel contrasts was measured for formant values (F1 & F2) for the four female subjects of each L1. Comparing the oral vowel spaces of Tamil English (TE) and Gujarati English (GE) monophthongs, by comparing the mean F1 and F2, reveals differences in these dialects due to transfer rather than similarities based on the single IE target system. For example, GE back vowels show very small spectral differences: [u], [upsilon], and [o] are grouped closely together within the vowel space; GE [u] is shorter in duration than the other two. The apparent lack of an /upsilon/-/o/ distinction in GE may be the result of transfer from Gujarati, which lacks a short high back vowel similar to /upsilon/ (Mistry 1997).

Consonantal differences between the dialects were determined through transcription, and several effects of L1 transfer were found. Shared phonetic features were realized differently. For example, both TE and GE use retroflex stops, following the General IE model; however, retroflexion was stronger for TE speakers than for GE speakers, likely due to transfer of L1 retroflexion norms (cf. Ladefoged and Bhaskararao 1983). Individual segments were also transferred, such as the Tamil fricative rhotic vs. the Gujarati tap. Allophonic variation also differed by L1 in the voiceless stop series. While Tamil has no phonemic contrast in voicing/aspiration, Gujarati has a 4-way contrast. Speakers of both varieties made a consistent distinction between voiced and voiceless stops; however, TE speakers aspirate initial stops, as they do in Tamil (Balasubramanian 1975), while GE speakers do not. Thus the consonantal differences shows transfer of phonetic features, segments, and allophonic variation from L1.

For the prosodic characteristics, a preliminary intonation analysis, based on a transcription (closely following the ToBI system for AE) of a set of read sentences (e.g., declarative, yes/no questions, wh-questions, focus) revealed systematic differences between TE and GE. In general, IE utterances appeared to have many pitch accents assigned to words prior to the boundary of the intonation phrase. In TE and GE, most or all content words (or sets of function words) prior to the edge of the intonation phrase are a pitch accent; GE speakers typically use a rising pitch accent that might be best transcribed as L*+H or L+H*, while TE speakers use either a falling pitch accent (H+L* or H*+L), a high pitch accent (H*), or a rising pitch accent (L*+H or L+H*). Tamil speakers used H+L* significantly more frequently than GE

speakers, while GE speakers used L*+H with significantly greater frequency. The use of multiple pitch accents in single utterances appears to be an important feature of IE intonational phonology, with TE and GE favoring distinct pitch accents (cf. Ravishankar 1994).

Thus, the data reveal both phonetic and phonological influences of L1 on the L2, even in proficient speakers; these influences appear to supercede IE norms and can be found in both the segmental and suprasegmental properties of their speech.

References cited

Balasubramanian, T. 1975. Aspiration of voiceless stops in Tamil and English -- An instrumental investigation. *CIEFL Newsletter*, 14-18.

Mistry, P.J. 1997. Gujarati Phonology. In *Phonologies of Asia and Africa*, Ed. by Alan S. Kaye, Eisenaous: Winona Lake, Indiana, 652-673.

Ravisankar, G. 1994. *Intonation Patterns in Tamil*. Pondicherry Institute of Linguistics and Culture: Pondicherry, India.

The Interlanguage Phonology of Mandarin Learners of English and the Gradual Learning Algorithm

Zheng Xu
State University of New York-Stony Brook

Model: This paper reports on an attempt to model the longitudinal development of Mandarin Chinese speakers learning English and variation in their production patterns using the Gradual Learning Algorithm (GLA) of Boersma & Hayes (2001). I assume that the learning process begins with the native language grammar, which changes in response to the target language input. The GLA predicts that the ways in which the grammar changes are directly determined by the frequency of various structures in the target language input. Additionally, the GLA relies on stochastic rankings of constraints allowing us to model a speaker's output variation.

Experimental data: I first carried out an experimental investigation of native Mandarin speakers pronouncing English words containing final obstruents and labial nasals, none of which occur in Mandarin codas. The experiment revealed that each subject employed different strategies in producing word-final obstruents and labial nasals. For example, the strategies in producing final voiced obstruents (e.g., [???]) include consonant deletion (e.g., [??]), vowel epenthesis (e.g., [??.??]), final devoicing (e.g., [???]), and correct pronunciation (e.g., [???]). The strategies in producing word-final [?] (e.g., [???]) include consonant deletion (e.g., [??]) and correct pronunciation (e.g., [???]).

However, certain generalizations emerged. First, all 8 subjects correctly pronounced more word-final voiceless obstruents than word-final voiced obstruents. Second, seven of the eight subjects correctly pronounced more word-final [?] than word-final voiced obstruents. And third, there was a rough correlation between experience in English and error patterns, with consonant deletion and vowel epenthesis decreasing and final devoicing increasing with longer EFL experience.

Simulation results: I compared the Mandarin learners' production patterns with the GLA simulation results. The GLA correctly predicts the development of coda obstruents, predicting an initial stage in which voiced and voiceless coda obstruents are deleted or become onsets through vowel epenthesis, followed by a stage in which voiceless obstruents are more often correctly produced than voiced ones. Additionally, the GLA correctly predicts that devoicing of final voiced obstruents becomes more frequent and vowel epenthesis after final voiced obstruents and deletion of final voiced obstruents become less frequent as a Mandarin speaker's EFL experience increases. Furthermore, the GLA correctly models the variation seen in the data between correct production, vowel epenthesis after final consonants, deletion of final consonants, and devoicing of final voiced obstruents.

However, the GLA did not correctly model the relative order of acquisition of obstruents and labial nasals. The GLA predicts that Mandarin learners of English will correctly produce more word-final voiced obstruents than [?], for word-final voiced obstruents are more abundant than coda labials in the English input, which should cause the markedness constraints on word-final voiced obstruents to be demoted more quickly than the markedness constraints on word-final labials. However, speakers made many fewer errors with final labial nasals than with final voiced obstruents. This suggests that other factors such as perception may also affect the L2

acquisition process and that the current GLA model needs improvement to incorporate these factors.

Early peak alignment and deep falls in Buenos Aires broad focus declaratives

Laura Colantoni & Jorge Gurlekian
University of Toronto and Laboratorio de Investigaciones sensoriales-Conicet

Previous autosegmental analyses of Spanish broad focus declarative utterances consistently report a late peak alignment for pre-nuclear pitch accents in broad focus declaratives (Sosa 1991, 1999; Prieto et al. 1995; Face 2001, 2002; Hualde 2000, 2002; Beckman et al. 2002, among others). We seek to demonstrate here that Buenos Aires Spanish presents a different alignment pattern for pre-nuclear accents, i.e. a peak aligned within the stressed syllable. An additional feature that has been proposed to characterize this Spanish variety is the presence of a long falling contour (Vidal de Battini 1964; Kaisse 2001). We tested these hypotheses through the analysis of a corpus of 1200 sentences recorded by a female and male speaker from Buenos Aires. Data were manually labeled twice by four research assistants, using a modified version of ToBI (Beckman and Ayers 1993) that associates an ERB value with each tonal event (Gurlekian et al. 2001). Results confirmed both hypotheses. First, a fundamental frequency (F0) peak is aligned within the stressed syllable in 90% of the instances in pre-nuclear pitch accents. Early peak alignments have been observed in Italian (Ladd 1996) in this same context, but have only been reported in other Spanish varieties in a specific pragmatic context, i.e. to convey contrastive focus (Face 2001, Beckman et al. 2002, among others). Second, final intonational phrases consist of a pronounced falling contour, which ends up near the speaker's tonal baseline (ERB values associated with final boundary tones are typically 2.5 ERBs lower than the initial boundary tone). We argue here that these differences emerged through (1) the loss of a pragmatic contrast in a particular sociolinguistic situation, and (2) contact with Italian. First, contrastive/emphatic patterns available in Spanish became the default accent markers in Buenos Aires Spanish, in a grammaticalization-like process. This hypothesis is supported both by historical data, and by research on focus. Regarding the former, Vidal de Battini (1964) comments that in the turn of the XX century there was a change in intonation patterns in Buenos Aires Spanish, which she attributes, in part, to an extension in the use of emphatic contours. Those contours were originally used by gang members and criminals, and spread to other social groups. On the other hand, research on focus (Toledo 1989) demonstrates that Buenos Aires Spanish does not consistently use early peak alignments to mark focus, as other varieties do. The absence of consistent tonal marks can be attributed to the fact that Buenos Aires Spanish already uses early peak alignments in broad focus declaratives. The second hypothesis that would explain the differences observed in Buenos Aires Spanish involves contact with Italian (Vidal de Battini 1964; Kaisse 2001). This hypothesis is supported by the existence of a large population (around 60%) of Italian descent. In addition, recent research on Italian intonation shows that pre-nuclear accents consist of a peak within the stressed syllable (Ladd 1996), and that declarative contours of some varieties (such as Neapolitan) are characterized by a pronounced falling contour (D'Imperio 2002). Thus, following Ladd's (1996) proposed typology, Buenos Aires Spanish and other varieties previously described would in principle have 'semantic differences', since identical tunes are used to convey different meanings. It is a future task to determine whether the additional differences are purely realizational or also systemic.

Reference List:

Beckman, Mary and Ayers Elam, Gayle. Guidelines for ToBI labeling. http://ling.ohio-state.edu/phonetics/E_ToBI . 1993. The Ohio State University Research Foundation.

Beckman, Mary, Díaz-Campos, Manuel, Tevis McGory, Julia, and Morgan, Terrell. 2002. Intonation across Spanish in the Tones and Break Indices framework. Probus 14.9-36.

D'Imperio, Mariapaola. 2002. Italian intonation: an overview and some questions. Probus 14.37-69.

Face, Timothy. 2001. Focus and early peak alignment in Spanish intonation. Probus 13.223-46.

Face, Timothy. 2002. Local intonational marking of Spanish contrastive focus. Probus 14.71-92.

Gurlekian, Jorge, Rodríguez.Hernán, Colantoni, Laura, and Torres, Humberto. Development of a prosodic database for an Argentine Spanish text to speech system. Proceedings of the Workshop on linguistic databases. 2001. Philadelphia, Uniersity of Pennsylvania.

Hualde, José Ignacio. 2000. Intonation in Spanish and other Ibero-Romance languages: Overview and status questionis, ed. by Wiltshire, Caroline and Josquin Camps, Amsterdam: Benjamins.

Hualde, José Ignacio. 2002. Intonation in Romance. Introduction to the special issue. Probus 14.1-7.

Kaisse, Ellen. 2001. The long fall: an intonational melody of Argentinean Spanish, ed. by Herschensohn, Julia, Enrique Mallen, and Karen Zagona, 148-60. Amsterdam: Benjamins.

Ladd, Robert. 1996. Intonational phonology. Cambridge: Cambridge University Press.

Prieto, Pilar, Van Santen, Jan, and Hirschberg, Julia. 1995. Tonal alignment patterns in Spanish. Journal of Phonetics 23.429-51.

Sosa, Juan Manuel. Fonética y fonología de la entonación del español hispanoamericano. 1991. University of Massachussets.

Sosa, Juan Manuel. 1999. La entonación del español: su estructura fónica, variabilidad y dialectología. Madrid: Cátedra.

Toledo, Guillermo. 1989. Señales prosódicas del foco. Revista argentina de lingüística 5.205-30.

Vidal de Battini, Berta. 1964. El español de la Argentina. Buenos Aires: Consejo Nacional de Educación.

French Intonational Structure: Evidence from Tonal Alignment

Pauline Welby
Institut de la Communication Parlée, CNRS UMR 5009 Institut National Polytechnique de Grenoble, Université Stendhal Grenoble

A production study examined the structure of French intonation, which is characterized by an obligatory fundamental frequency (F0) rise on the last syllable of a non-utterance-final phrase (1a) and an optional early rise before the late rise (1b). (There is no meaning difference.)

(1) a. [Les gamins SAGES] jardinaient. 'The good kids were gardening.'

(2) b. [Les GAmins SAGES] jardinaient.

The 2-rise pattern is often treated as a sequence of L(ow) and H(igh) tones: LHLH ([2,3,4,5,6]).

In keeping with the theme of the conference, we note that the early rise is a relatively recent development in the language. While it is quite common, accepted, and completely unremarkable in all registers of present-day French, as late as the 1950s, it was berated as an undesirable element that threatened the "beauty of the French language" (See discussion in [3]).

A set of sentences was designed to examine the alignment of both rises and the factors influencing the appearance of the early rise. Seven speakers read the corpus at two speaking rates, producing 790 usable utterances. Word and syllable boundaries and F0 peaks and valleys were labelled for each target phrase. Time, duration, and F0 values were extracted from the label files.

The two-rise LHLH pattern was most common. Duration in clock time (ms) and length in syllables were equally good predictors of this pattern. Length in syllables (but not duration) has been proposed as a predictor by some models. Several other patterns were also observed: LH, LLH, LHH, LHi, and L2H. The LHi is a pattern with an early rise, but no late rise---its use is appropriate for an enumeration or list. One of the patterns, the L2H, is not reported in the literature. In fact, Jun and Fougeron (2002) predict that this pattern should be impossible. The pattern, rare even in our data, is characterized by a fall from the late H of a preceding phrase to a L in the last syllable of the target phrase, followed by a rise to the late H.

An analysis of the text-to-tune alignment, the timing of F0 peaks and valleys with respect to segmental landmarks, showed that the two tones at the edges of the LHLH pattern were most stably aligned. The early L was consistently aligned with the function word/content word boundary. The late H was consistently realized near the end of the last syllable of the phrase. There was much more variation in the timing of the H of the early rise and the L of the late rise. In addition, the H of the early rise was often realized as a high plateau. The alignment facts are inconsistent with the Di Cristo and Hirst model in which the early and late rises have the same structure. In addition, unlike the H of the late rise, the H of the early rise shows no evidence of association to a syllable, evidence against the claim in [6] that it is a pitch accent.

We follow Jun and Fougeron (2002) in treating the late rise as a LH bitonal pitch accent and the early rise as a LH bitonal phrase accent. We argue that the early L of the early rise is "edge-seeking"-- it is doubly associated to the left edge of the content word and to the left edge of an earlier syllable boundary (a revision of the claim in [7]). We note that standard assumptions of Autosegmental Metrical theory cannot account for the timing of the interior tones of the LHLH pattern, which behave neither like associated tones nor like trailing/leading tones. Similar observations have been made for Greek ([1]).

The results of the production study guided the development of a series of perception experiments on the role of intonational cues in speech segmentation ([8,9]). The production data revealed an alignment pattern unreported in the literature: Low F0 inflections or "elbows" were sometimes realized at function word/content boundaries, even when there was no rise (no following H). The perception results showed that listeners use both an early rise and a simple F0 elbow as cues to content word beginnings. The F0 elbow finding illustrates the value of developing perception studies guided by the results of complementary production studies.

References

[1] Arvaniti, Amalia, D. Robert Ladd, and Ineke Mennen. (2000). What is a starred tone? Evidence from Greek. In M.B. Broe and J.B. Pierrehumbert, eds. Papers in laboratory phonology V: acquisition and the lexicon. Cambridge: CUP.

[2] Di Cristo, Albert and Hirst, Daniel. (1996).Y-a-t-il des unités tonales en français ? 20th J.E.P. pp. 223-226. Avignon.

[3] Di Cristo, Albert. (1999). Vers une modélisation de l'accentuation du français: première partie. French Language Studies 9: 143—179.

[4] Di Cristo, Albert. (2000). Vers une modélisation de l'accentuation du français: seconde partie. French Language Studies 10: 27—44.

[5] Jun, Sun-Ah and Cécile Fougeron. (2002). Realizations of Accentual Phrase in French. Probus 14: 147—172. [A special issue on the intonation of Romance languages, edited by J. Hualde.

[6] Post, Brechtje. (2000). Tonal and phrasal structures in French intonation. The Hague: Thesus.

[7] Welby, Pauline. (2002). The realization of early and late rises in French: a production study. In B. Bel and I. Marlien, eds., Proceedings of the Speech Prosody 2002 conference, pp. 695-698. Aix-en-Provence.

[8] Welby, Pauline. (2003a). French Intonational Rises and their Role in Speech Seg Mentation. In Proceedings of Eurospeech: the 8th annual conference on speech communication and technology, pp. 2125—2128. Geneva.

[9] Welby, Pauline. (in preparation). The slaying of Lady Mondegreen, being a study of French tonal association and alignment their role in speech segmentation. PhD dissertation, The Ohio State University.

Learning different phonetic distributions in onset and coda position

Kyle E. Chambers*, Kristine H. Onishi*, Cynthia Fisher* & Jessica Maye†
University of Illinois and University of Rochester†*

How do phonological representations develop and change? It is sometimes assumed that infants learn which contrasts matter in their native language by noting semantic minimal pairs--the semantic contrast between "rake" and "lake" requires that the distinction between 'r' and 'l' is contrastive in English. However, recent findings suggest that both adults' and infants' tendency to discriminate a phonetic contrast can be affected by distributional properties of their experience with that contrast, without support from semantic information (Maye & Gerken, 2001; Maye, Werker, & Gerken, 2002). Training with a bimodal distribution of values along a phonetic continuum (i.e., values near both extremes of the phonetic continuum, with few values at the center) suggests the existence of two categories, and increased discrimination of the endpoints of the continuum in a later test. Training with a unimodal distribution (i.e., values clustering toward the middle of the continuum, with few values at the extremes) suggests a single category, and decreased discrimination of the endpoints.

However, a bimodal distribution of phonetic values does not always correspond to a phonemic contrast for speakers of that language. The key to determining whether a phonetic difference is contrastive or allophonic is whether values on the relevant phonetic dimension are bimodally distributed in the _same_ phonological environment (e.g., Peperkamp & Dupoux, 2003).

The current study explores whether adults implicitly track both the distribution of values on a phonetic dimension and the phonological environment of that distribution. We chose a voiced and an unvoiced uvular fricative (denoted "gh" and "x", respectively), produced by a native speaker of Arabic, as a contrast that would be unfamiliar to English speakers. From naturally spoken consonant-vowel-consonant syllables containing the uvular fricatives, we synthesized multiple 8-step continua of syllables ranging from one endpoint containing a voiced fricative to another containing an unvoiced fricative in the same syllable position (e.g., "ghik" to "xik" and "bigh" to "bix"). The syllables from these continua were presented with varying frequency of occurrence to monolingual English-speaking subjects. Half of the subjects heard the uvular fricative with a bimodal distribution along the voicing dimension in onset position, but a unimodal distribution in coda position. The other subjects heard the uvular fricative with a unimodal distribution in onset position, but a bimodal distribution in coda position. Later, subjects were asked to judge whether pairs of these stimuli, such as "ghik" and "xik," and "bigh" and "bix," sounded like different words or two utterances of the same word. If subjects separately kept track of the phonetic distributions in different phonological environments (onset and coda), then for either onset or coda position, subjects who received the bimodal training in that syllable position should respond "different" more often than those who received unimodal training in the same position. Preliminary results confirm this prediction.

This pattern suggests that adults are able to acquire and maintain separate distributions of phonetic values for different phonological environments. Furthermore, they are able to do this with a nonnative consonant contrast. These findings lend further support to the important role of distributional learning in phonological development; this work may ultimately help to explain how language learners uniformly converge on a phonological inventory for their native language before amassing a very large vocabulary of meaningful words.

An investigation of front-back confusions in vowels

Jose R. Benki
University of Michigan

One of the most documented trends in diachronic vowel changes is for the back high and mid vowels (/u o/) to move forward in the vowel space (Sweet, 1888). This trend is given axiomatic status by Labov (1994) as his third principle in vowel chain shifts: back vowels move to the front. The present study investigates the perceptual and articulatory basis for this trend.

Previous research (Benki, 2003) indicates that for English nonsense monosyllables presented in background noise, the front-back dimension, conveyed primarily by the second formant (F2) is the most vulnerable, with most vowel confusions between front and back vowels of the same height. The observed pattern of confusions was asymmetric in that the back vowels /u/ and /o/ were perceived as the front vowels /i/ and /e/ respectively much more than the vowels /i/ and /e/ were perceived as /u/ and /o/. Both the vulnerability of the front-back dimension as well as the asymmetry of the confusions are consistent with older work reporting confusions for vowels in noise (Pickett, 1957; Nooteboom, 1968) as well as the historical tendency of back vowels to move forward.

Mechanisms for the perceptual vulnerability of the front-back dimension and robustness of the height dimension in vowels have been documented (Lindblom, 1986) and would explain the confusions in the front-back dimension. Neither the asymmetry of the reported experimental data nor the observed direction of vowel shifts are not easily explained by the same mechanism, however.

An articulatory explanation for the asymmetry could be the tendency of back vowels, particularly /u/, to be slightly fronted in alveolar and velar contexts, while front vowels, particularly /i/, are highly to backing (Recasens, 1999). In noise, listeners might fail to compensate for the coarticulation and misperceive (slightly) fronted back vowels as front vowels. The articulatory explanation predicts that the asymmetry should be reduced in an experiment with hyperarticulated back vowels (i.e., without fronting) as stimuli. Experiment 1 tests this prediction by evaluating listener confusions of CVC stimuli in noise where the vowels are /i e u o/ and fully-backed /u o/, and both consonants are either bilabials, alveolars, or velars.

A two-part perceptual explanation for the asymmetry is proposed, involving the noise-induced spread of phase-locking (neural synchrony) with F1 to other nerve fibers with low characteristic frequencies (CFs) that would otherwise synchronize to F2 for a back vowel. Simultaneously, the masking noises induce the percept of energy consistent with F2 of a front vowel (Shriberg, 1992). This explanation predicts that asymmetrical confusions should be reduced for stimuli with the spectral information replaced with band-limited noise at different frequency bands, with each band temporally modulated by the original signal. moved (thus reducing neural synchrony) and replaced with noise in separate frequency bands modulated according to the temporal envelopes of the appropriate frequencies in the original stimuli, following Shannon et al. (1995). Experiment 2 tests this prediction by evaluating listener confusions of stimuli from Experiment 1, transformed following Shannon et al. (1995).

References

Benki, J. R. (2003) Analysis of English nonsense syllable recognition in noise, Phonetica 60, 129-157.

Lindblom, B. (1986) Phonetic universals in vowel systems; in Ohala, Jaeger, Experimental phonology (Academic Press, New York).

Nooteboom, S. (1968) Perceptual confusions among Dutch vowels presented in noise; in IPO Annual Progress Report, vol. 3, pp. 68-71 (Instituut voor Perceptie Onderzoek, Eindhoven).

Pickett, J. M. (1957) Perception of vowels heard in noises of various spectra. J. acoust. Soc. Am. *29,* 613-620.

Recasens, D. (1999) Lingual coarticulation; in Hardcastle, Hewlett, Coarticulation: theory, data and techniques (Cambridge University Press, Cambridge).

Shannon, R. et al. (1995) Speech recognition with primarily temporal cues, *Science* 270, 303-304.

Shriberg, E. E. (1992) Perceptual restoration of filtered vowels with added noise, Language and Speech 35, 127-136.

Sweet, H. (1888) History of English sounds. Clarendon Press, Oxford 1888.

Nonlinear links between continuity and discreteness: transparency in vowel harmony

Stefan Benus*, Adamantios Gafos* & Louis Goldstein[†‡]
New York University, Yale University[†] and Haskins Laboratories[‡]*

A central question in the background of much current work in laboratory phonology is how to relate the continuous dimensions of phonetics to categorical phonological alternations (Pierrehumbert et al. 00). We address an instance of this question from transparency in vowel harmony. We relate our experimental results to the vowel mergers leading to transparency, assuming that diachronic change is shaped by perceptual and articulatory factors alike. We then propose a model of the relation between the observed properties of transparent vowels and their phonology, using basic tools of nonlinear dynamics.

In Hungarian harmony, certain front vowels may intervene between the trigger and target vowels of the harmony even if they bear the opposite value for the harmonizing feature: in kávé-nak 'coffee' the first ([+back]) vowel dictates the backness for the suffix across the so-called transparent [é], which is [–back] (accent denotes length). Using EMMA (Perkell at al. 92), we compared the location of receiver attached to the rear of the tongue dorsum during transparent vowels (TVs) in pairs like zefír-nek 'zephyr', zafír-nak 'sapphire', with matched consonants (22 pairs, 16 repetitions). For all three subjects, TVs in stems selecting back suffixes are articulated further back than TVs in stems selecting front suffixes, e.g. for one subject the mean difference for the tongue dorsum receiver in the two contexts is .4mm ($p<0.001$). A similar result obtains for the so-called abstract stems. Monosyllabic stems with TVs usually select front suffixes as in víz 'water(.NOM)', víz-nek. A limited number (about fifty) of these stems, however, select back suffixes as in híd 'bridge(.NOM)', híd-nak. The tongue dorsum for phonemically identical transparent vowels in such stems produced in isolation (no overt suffix) is more retracted in the stems selecting back suffixes than in the stems selecting front suffixes (mean difference = .5mm, $p<0.01$). The overarching generalization then is that non-distinctive retraction in TVs, either lexically-specified as in the monosyllables or contextually determined as in the cases of zefír-/zafír-, correlates with the phonological alternation in suffix form; the advanced/retracted version of a TV selects the front/back suffix. The diachronic evidence available reasonably suggests that the change leading to transparency was perceptually based, originating from loss of contrast or merger between front and unround back vowels in prosodically weak contexts (Kálmán 72). Our data, however, shows that there is a discernible distinction in articulation. We propose that the incompleteness of the merger is due to the articulatory pressures of agreement imposed by harmony, e.g. [í] in zafír- is retracted due to an agreement relation with its preceding vowel.

How then are fine differences in articulation to be linked to the phonological alternations of vowel harmony? Formally, the relation between retraction in the TV and suffix form is nonlinear: small changes in degree of retraction are associated with large changes in suffix form. In the proposed dynamical model, the two discrete forms of a suffix, say FRONT /–*nek*/ vs. BACK /–*nak*/ dative, are mapped to attractors of a dynamical system. Building on Gafos (in press), this is modeled by a function with two minima at the values of constriction location corresponding to FRONT /–*e*/, BACK /–*a*/. We require that the choice of the attractor, the surface form of the suffix, be determined by variation in R, representing the degree of tongue dorsum retraction for the stem-final vowel. Mathematically, these ideas can be stated succinctly in the form of the equation $dx/dt = N(x,R) + F(t)$, whose solutions correspond to the attractors of constriction location for the suffix (N is a nonlinear function over x, R, and F(t) represents random noise).

We show that for a range of small R values, the system is mono-stable with one attractor in the front region of the state space. As R is increased beyond a certain critical value, there is a change from the FRONT to the BACK attractor. In this model, then, a discrete change in suffix form is brought about by a scalar increase in the continuous retraction degree R of the transparent vowel. Effectively, non-contrastive differences in R can result in categorical alternations. This is due to the fundamental property of nonlinearity which allows us to link changes along continuous dimensions to categorical alternations.

References

Gafos, A. In press. Dynamics in Grammar: Comment on Ladd and Ernestus and Baayen. In Louis M. Goldstein, D. H. Whalen, Catherine T. Best, and Stephen R. Anderson (eds.), Varieties of Phonological Competence, Berlin/New York: Mouton de Gruyter.

Kálmán, B. 1972. Hungarian historical phonology. In L. Benko and S. Imre (eds.), The Hungarian Language, The Hague: Mouton, 49-83.

Perkell, J.S., Cohen, M., Svirsky, M., Matthies, M., Garabieta, I. and Jackson, M. 1992. Electro-magnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements. Journal of the Acoustical Society of America 92, 3078-3096.

Pierrehumbert, J., Beckman, M. E and Ladd, D. R. 2001. Conceptual Foundations of Phonology as a Laboratory Science. In Burton-Roberts, N., Carr, P. and Docherty, G. (eds.), Phonological Knowledge, Oxford, UK: Oxford University Press, 273-304.

Vowel phonotactics and syllabification in native and non-native phonology

Juli Cebrian
Universitat Autonoma de Barcelona

In addition to learning the segmental structure, L1 language learners come to know the syllabic and prosodic structure and the phonotactic constraints of their language (Lindblom, 1990; Strange, 1999). Research indicates that phonotactic constraints are a distinct component of phonological knowledge and that infants, children, and even adults have a predisposition to extract these patterns from speech (Vitevich and Luce, 1998; Onishi, Chambers and Fisher, 2002; Mattys and Jusczyck, 2001). It seems reasonable to assume that phonotactics and syllabification are learned concomitantly with the other components of the speaker's L1 grammar and thus it is possible that they are acquired in the course of the acquisition of a second language. However, as with segmental structure, the acquisition of L2 suprasegmentals may be affected by interference from the L1. Studies have in fact shown that L1 syllabification and phonotactic constraints are indeed carried over to the L2 and that interference from L1 suprasegmental structure may influence the pronunciation of L2 sounds (Tarone, 1987; Trammell, 1993; Rochet and Putnam Rochet, 1999, Delattre, 1965; Solé, 1989).

This study focuses on the English phonotactic contraint that disallows lax vowels in stressed open syllables, that is, the Lax Vowel Constraint (LVC), and on syllabification patterns that may be triggered by this constraint. A number of studies on syllabification strategies in English have found that coda syllabification of intervocalic consonants is more common with preceding lax vowels than with tense vowels (Treiman and Danis, 1988; Treiman and Zukowski, 1990; Derwing, 1992; Fallows, 1981). This paper thus examines the prevalence of the LVC and syllabification patterns in native English speakers and also investigates whether L2 learners of English possess knowledge of the LVC and if it has any influence in the way they structure English words. A group of 25 native speakers of Canadian English and a group of 30 adult learners of English and native speakers of Catalan participated in two singular-plural picture naming experiments and two syllable manipulation experiments aimed at assessing knowledge of the LVC and syllabification preferences of the two groups of speakers.

The results for native English speakers were consistent with previous research that shows a preference for coda syllabification of medial consonants when the preceding vowel is lax. With respect to picture naming tasks, native speakers yielded a very high percentage of LVC-obeying responses, that is, disallowing lax vowels in open syllables. This supports the view of the LVC as an active constraint for native English speakers, which determines their knowledge of syllable and word well-formedness. The results for non-native speakers showed that the L2 learners have some knowledge of the phonotactic restriction on lax vowels. Although it is not as strong as with native speakers, Catalans do show a tendency to avoid lax vowels in open syllables. This is clearly not an L1 feature, and it provides evidence that the LVC is actually learned. Still, the L2 learners showed a preference for onset syllabification of medial consonants, illustrating a strong influence of L1 syllabification strategies, consistent with previous L2 acquisition studies. Taken together, the results for both native and non-native speakers show that the LVC is part of the phonological system of both L1 and L2 English. The LVC is fully established in L1 English but less strongly manifested with L2 learners, resembling the way in which native speakers perceive and produce vowels more accurately than L2 learners.

References

Delattre, P. 1965. Comparing the phonetic features of English, French, German and Spanish. New York: Chilton.

Derwing, B. 1992. A 'pause-break' task for eliciting syllable boundary judgements from literate and illiterate speakers: preliminary results for five diverse languages, Language and Speech, 35, 219-235.

Fallows, D. 1981. Experimental evidence for English syllabification and syllable structure. Journal of Linguistics, 17, 309-318.

Lindblom, B. 1990. Explaining phonetic variation: A sketch of the HandH theory. In W. J. Hardcastle. & S. Marchal (Eds.), Speech production and speech modeling. Dordrecht: Kluwer Academic.

Mattys, S.L. & P.W. Jusczyk. 2001. Phonotactic cues for segmentation of fluent speech by infants. Cognition, 78, 2, 91-121.

Onishi, K.H., K.E. Chambers & C. Fisher. 2002. Learning phonotactic constraints from brief auditory experience. Cognition 83, B13-B23.

Rochet, B.L. & A. Putnam Rochet. 1999. Effects of L1 phonotactic constraints on L2 speech perception. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A. Baily (Eds.) Proceedings of the 14th International Congress of Phonetic Sciences (pp. 1443-6). Berkeley, CA: University of California.

Solé, M.J. 1989. Phonotactic constraints as a filter of the articulatory-acoustic signal. Journal of the Acoustical Society of America Supplement 1, Vol. 86, S18.

Strange, W. 1999. Levels of abstraction in characterizing cross-language phonetic similarity. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A. Baily (Eds.) Proceedings of the International Congress of the Phonetic Sciences, San Francisco, CA, 1999. (pp. 2513-9).

Tarone, E. 1987. Some influences on the syllable structure of interlanguage phonology. In G. Ioup & S. H. Weinberger (Eds.), Interlanguage phonology: The acquisition of a second language sound system (pp 232-247). New York: Newbury House.

Trammell, R.L. 1993. English ambisyllabic consonants and half-closed syllables in language teaching. Language Learning, 43, 195-238.

Treiman, R. & C. Danis. 1988. Syllabification of intervocalic consonants, Journal of Memory and Language, 27, 87-104.

Treiman, R. & A. Zukowski. 1990. Toward an understanding of English syllabification, Journal of Memory and Language, 29, 66-85.

Vitevich, M.S. & P.A. Luce. 1998. When words compete: Levels of processing in perception of spoken words. Psychological Science, 9, 325-329.

Markedness from the grammar, not the lexicon

Andries W. Coetzee
University of Massachusetts and Potchefstroom University

The locus of markedness effects has been the focus of research in recent years. Are so-called markedness effects epiphenomena of lexical statistics, or do they originate in the grammar? Some studies have shown that lexical statistics can account for relative well-formedness judgments, patterns in phoneme identification, and response times in lexical decision (e.g., Newman et al., 1997, Pierrehumbert et al., in press, Vitevitch & Luce, 1998). Others have shown that these effects can also sometimes be accounted for by relying on grammar as such (Berent & Shimron, 1997, Frisch & Zawaydeh, 2001, Moreton, 2002).

In this paper, I report on two experiments showing that there are markedness effects that cannot be attributed to lexical statistics. This shows that markedness does originate in the grammar. It also shows that phonological grammar does influence linguistic performance directly.

Experimental tokens. English allows words of the form [sTvT] (*state, stout*), but not of the form [sKvK] or [sPvP] (**skake,* **spape*) (Davis, 1991, Fudge, 1969). There are also languages, such as Afrikaans, that allow words of the form [sTvT] and [sKvK] (*staat* "state", *skok* "shock"), but not of the form [sPvP] (**spaap,* **spop*). Based on these facts, I propose the following universal markedness scale: sTvT > sKvK > sPvP.

Experiment 1: Comparative well-formedness. I selected fifteen non-word pairs of the form [sTvT]~[sKvK]. In each pair the cumulative biphone probability was higher for the [sKvK]- than the [sTvT]-member, and for twelve of the pairs the lexical neighborhood density was higher for the [sKvK]- than the [sTvT]-member. Lexical statistics therefore favored the more marked member of a pair. These pairs were presented to subjects who had to select the member in a pair that was most word-like. In spite of the fact that lexical statistics favored the more marked [sKvK]-member, subjects consistently preferred the less marked [sTvT]-member ($t(514)=12.133$, $p<.000$). A similar comparison was done with [sTvT]~[sPvP]-pairs ($t(509)=14.669$, $p<.000$) and [sKvK]~[sPvP]-pairs ($t(502)=2.737$, $p<.003$). In all three conditions, the less marked member was preferred in spite of the fact that the lexical statistics favored the more marked member.

Experiment 2: Lexical decision. I selected five non-words of the form [sTvT] and five of the form [sKvK]. The tokens were selected such the lexical neighborhood densities and cumulative biphone probabilities of the tokens in the two groups did not differ significantly. These tokens were presented in a speeded lexical decision task to subjects. Reaction times for correct non-word responses were slower for the less marked [sTvT]- than the more marked [sKvK]-tokens ($t(375)= 2.97$, $p < .002$). A similar comparison was done between [sTvT]- and [sPvP]-forms ($t(366)= 3.07$, $p < .001$), and between [sKvK]- and [sPvP]-forms ($t(371) = 1.90$, $p < .025$). The general result was that reaction times for the different groups were related as follows: sTvT > sKvK > sPvP. Since the lexical neighborhood densities and cumulative biphone probabilities of the tokens were controlled for, this is attributed to the relative markedness of the three classes.

Discussion. In *Experiment 1*, the results are the opposite from what would have been expected based on the lexical statistics. However, the results follow the universal markedness scale. In *Experiment 2*, the tokens did not differ in terms of their lexical statistics, so that lexical

statistics cannot be responsible for patterns observed in the results. The results again follow the universal markedness scale. I interpret these results as showing that markedness is located in the grammar, and not merely a reflex of lexical statistics. The difference that was found between [sKvK]- and [sPvP]-forms is particularly strong evidence to locate markedness in Universal Grammar. Neither of these kinds of words is allowed in English, and even so English speakers show a preference for [sKvK]-forms over [sPvP]-forms.

References

Berent, Iris & Shimron, Jospeh. 1997. The representation of Hebrew words: evidence from the obligatory contour principle. Cognition, 64:39-72.

Davis, Stuart. 1991. Coronals and the phonotactics of nonadjacent consonants in English. In The Special Status of Coronals: Internal and External Evidence, eds. C. Paradis and J.F. Prunet, 49-60. San Diego: Academic Press.

Frisch, Stefan A. & Zawaydeh, Bushra Adnan. 2001. The psychological reality of OCP-Place in Arabic. Language, 77:91-106.

Fudge, E.C. 1969. Syllables. Journal of Linguistics, 5:253-286.

Moreton, Elliott. 2002. Structural constraints in the perception of English stop-sonorant clusters. Cognition, 84:55-71.

Newman, Rochelle S., Sawusch, James R. & Luce, Paul A. 1997. Lexical neighborhood effects in phonetic processing. Journal of Experimental Psychology, 23:873-889.

Pierrehumbert, Janet, Hay, Jennifer & Beckman, Mary. In press. Speech perception, well-formedness, and the statistics of the lexicon. In Papers in Laboratory Phonology VI, eds. J. Local, R. Ogden and R. Temple. Cambridge: Cambridge University Press.

Vitevitch, M.S. & Luce, Paul A. 1998. When words compete: levels of processing in spoken word recognition. Psychological Science, 9:325-329.

Production of English Intervocalic /d/ by Catalan Speakers in Derived and Non-derived Words

Susana Cortes, Mitsuhiko Ota & Alice Turk
University of Edinburgh

While L2 phonology is known to be influenced by the rules and processes of the learner's L1, not all L1 patterns transfer to the L2 to the same extent. Some studies indicate, for instance, that postlexical processes of the L1 are more likely to affect the L2 than do lexical alternations (Altenberg & Vago, 1983; Weinberger, 1994; Zsiga, 1995). Moreover, effects of postlexical processes do not appear equally in all L2 contexts that satisfy the conditions of the application (Major & Faudree, 1996; Cebrián, 2000). In an attempt to characterize the L2 contexts that are more prone to L1 transfer, Eckman and Iverson (1997; also Eckman, Elreyes & Iverson, 2003) hypothesized that L1 postlexical rules can target (i) both derived and non-derived contexts or (ii) only derived contexts, but crucially not non-derived contexts only. In support of their hypothesis, none of the Spanish-speaking learners of English in their study showed a higher rate of intervocalic spirantization in non-derived English words (e.g., ladder) than in derived words (e.g., madder). The purpose of this paper is to further test this hypothesis by looking at the production of intervocalic /d/ by Catalan speakers of English.

Catalan has a spirantization rule which makes /d/ surface as a voiced interdental fricative in intervocalic position. Therefore, the Catalan word 'dia' (day) is pronounced with an initial stop whereas the word 'cada' (each) is pronounced with a fricative between the vowels. If Eckman and Iverson's hypothesis is right, then Catalan speakers will apply the L1 postlexical rule and, consequently, pronounce the intervocalic /d/ as a fricative either (i) in both derived and non-derived contexts (i.e. in madder and ladder) or (ii) only in derived contexts (i.e. in madder). If they have managed to eliminate the rule from their interlanguage, then they will pronounce intervocalic /d/ correctly in any context.

The production of English intervocalic /d/ by 20 native Catalan speakers with an advanced English level was elicited by asking them to read words containing intervocalic /d/ and to carry out a fill-in-the-blanks task to obtain intervocalic /d/ in derived words (e.g. lead-er). In the derived words, the originally final stop might become a fricative due to its becoming an intervocalic (i.e. fricative in Catalan) /d/. The data were auditorily analyzed by two phonetically trained native English speakers (MC and ZB). They were asked to identify the segment they heard in intervocalic position as a voiced alveolar stop or a voiced interdental fricative. If they thought the sound they perceived corresponded to none of the segments in the pair, they were to provide their own transcription.

Separate one-factor within subjects ANOVAs were run on the data transcribed by each analyzer. The production of intervocalic /d/ by Catalan speakers proved to be significantly better in non-derived than in derived words. This suggests that speakers follow the behaviour predicted by Eckman and Iverson. That is, it is the case that Catalan speakers apply the postlexical rules to derived words more than to lexical entries. However, we should stress the fact that, even though these Catalan speakers pronounce intervocalic /d/ better in non-derived than in derived contexts, their overall production of non-derived /d/ was not excellent. Their production of intervocalic /d/ in non-derived words was 45.61% correct in MC's analysis whereas it was 60.43% correct in ZB's analysis. Since this is not what we expected, we have collected some surveys to test

whether it is the frequency rather than the morphological configuration of words which affects their pronunciation.

Our findings suggest that Catalan speakers apply their L1 postlexical rule more in derived than in non-derived words in English. This shows that Catalan speakers might be using their L1 allophonic rule at the lexical level when they speak English and, therefore, they pronounce intervocalic /d/ as a fricative in derived English words before managing to eliminate the rule from their interlanguage.

References

Altenberg, E.P. & Vago, R.M. (1983). Theoretical Implications of an Error Analysis of Second Language Phonology Production. Language Learning, 33 (4), 427-447.

Cebrián, J. (2000). Transferability and Productivity of L1 Rules in Catalan-English Interlanguage. Studies in Second Language Acquisition, 22 (1), 1-26.

Eckman, F.R. & Iverson, G.K. (1997). Structure Preservation in Interlanguage Phonology. In S.J. Hannahs & M. Young-Scholten (Eds.), Focus on Phonological Acquisition. Vol.16. Collection: Language Acquisition & Language Disorders. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Eckman, F.R., Elreyes, A. & Iverson, G.K. (2003). Some principles of second language phonology. Second Language Research, 19 (3), 169-208.

Major, R. & Faudree, M. (1996). Markedness Universals and the Acquisition of Voicing Contrasts by Korean Speakers of English. Studies in Second Language Acquisition, 18, 69-90.

Weinberger, S.H. (1994). Functional and Phonetic Constraints on Second Language Phonology. In Mehmet S. Yavas (Ed.), First and Second Language Phonology. San Diego, CA: Singular Publishing Group Inc.

Zsiga, E.C. (1995). Phonology and Phonetics in the Education of Second Language Teachers: The Representation of Some Variable Rules of English. In J.E. Alatis, C. Straehle, B. Gallenberger, and M. Ronkin (Eds.), Proceedings of the 1995 Georgetown University Roundtable on Languages and Linguistics. Washington, DC: Georgetown University Press.

Fitting phonotactics into gestural phonology: Evidence from non-native cluster production

Lisa Davidson
New York University

Recently, it has been argued that syllable structure is reflected in the coordination of dynamical gestures (e.g. Browman and Goldstein 1995, Byrd 1996, Kochetov to appear). More specifically, these studies have focused on the nature of the coordination relationships that exist between gestures within the same syllable, but they do not discuss how syllabic affiliations are determined nor how languages impose restrictions on syllable formation. A gestural theory of syllable structure requires a mechanism for determining when multiple gestures can be combined into a syllabic constituent, like an onset. In other words, the mechanism for organizing gestures into syllables must interact with phonotactic restrictions. In this study, the representation of phonotactic restrictions in a gestural framework is developed using data from English speakers producing a series of English-illegal initial clusters. Experimental results show that clusters that differ along several dimensions are not produced with equal accuracy by the speakers even though none of them are grammatical in English. This indicates that non-native production can reveal aspects of L1 phonotactics that are not obvious from the L1 phonology alone.

The effects of voice, place, and manner in the production of non-native initial clusters were tested by combining fricative C1s with obstruent (O) and nasal (N) C2s to create clusters legal in Czech but not English. Experimental clusters included (1) /fm, fn, fs, fp, ft, fk/, (2) /zm, zn, zv, zb, zd, zg/, and (3) /vm, vn, vz, vb, vd, vg/, with (4) /sm, sn, sf, sp, st, sk/ as a control. 20 English speakers produced 4 pseudo-Czech words (e.g. *ftakay*, *vnaza*) for each onset cluster. The stimuli were presented both aurally, spoken by a native speaker of Czech, and in English-like orthography. The target cluster was repeated most accurately (i.e. like the Czech token) when C1=/s/ (97% correct), followed by /f/ (60%), /z/ (39%), and /v/ (22%). An ANOVA confirmed that all differences were significant ($p<.001$). Results also showed an effect of C2; for each C1, accuracy was significantly higher when C2=/nasal/ (65%) than /fricative/ (49%) or /stop/ (49%).

To incorporate these results into a gestural framework, the role of both perceptual and articulatory factors in defining phonotactics must be considered. Decreased accuracy for C1=/f, v/ relative to the coronals is perceptually motivated: low-intensity fricatives like /f/ and /v/ may be insufficiently detectable at cluster edges (Kingston 1990), so they are rare in that position. However, given that acoustic studies have shown that /z/ is significantly more perceptible than /f/ (Jongman, Wayland and Wong 2000), higher accuracy for C1=/f/ than C1=/z/ cannot be perceptual in origin. Instead, voicing in obstruent onset clusters may be difficult to maintain due to oral pressure buildup between the constriction and the glottis, which conflicts with the low pressure necessary to sustain voicing (Ohala and Kawasaki-Fukumori 1997). For C2, obstruents preferentially release into sonorants to maximize the perceptual cues for C1 (Steriade 1997).

Despite the multiple origins of the factors influencing phonotactic restrictions, the articulatory and perceptual features necessary for phonotactics and the temporal aspect of syllable structure can be merged at the phonological level. Within a constraint-based framework that includes temporal elements (Gafos 2002), the experimental clusters can be distinguished by a family of *Overlap/F?,F? constraints, where F? and F? refer to perceptual and articulatory features. *Overlap prevents gestures from having an overlapping coordination pattern if the gestures contain perceptual and/or articulatory features that cannot be combined into a cluster for that language. The fact that English speakers are not equally accurate on the experimental

clusters suggests that illegal clusters are governed by unique *Overlap constraints which are crucially ranked in the native grammar. This ranking is revealed when the production of non-native sequences forces a reassessment of what is treated as phonotactically legal. This proposal sheds light both on how native phonotactics may change through interaction with non-native sequences, and on how to incorporate phonotactics into a gestural account of syllable effects.

References

Browman, Catherine and Louis Goldstein. 1995. Gestural syllable position effects in American English. In Fredericka Bell-Berti and Lawrence Raphael, Eds., Producing Speech: Contemporary Issues for Katherine Safford Harris. New York: American Insititue of Physics.

Byrd, Dani. 1996. A phase window framework for articulatory timing. Phonology 13, 139-169.

Gafos, Adamantios. 2002. A grammar of gestural coordination. Natural Language and Linguistic Theory 20 (2), 269-337.

Jongman, Allard, Ratree Wayland and Serena Wong. 2000. Acoustic characteristics of English fricatives. Journal of the Acoustical Society of America 108 (3), 1252-1263.

Kingston, John. 1990. Articulatory binding. In John Kingston and Mary Beckman, Eds., Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech. Cambridge: Cambridge University Press.

Kochetov, Alexei. to appear. Syllable position effects and gestural organization: Articulatory evidence from Russian. In Papers in Laboratory Phonology VIII. Cambridge: Cambridge University Press.

Ohala, John and Haruko Kawasaki-Fukumori. 1997. Alternatives to the sonority hierarchy for explaining segmental sequential constraints. In Stig Eliasson and Ernst Hakon Jahr, Eds., Language and Its Ecology. Berlin: Mouton de Gruyter.

Steriade, Donca. 1997. Phonetics in Phonology: The Case of Laryngeal Neutralization. ms., UCLA.

Variability in Phonetic Implementation: Evidence from Korean Speakers' Perception of the Stop Epenthesis in English

Jeong-Im Han
Konkuk University

A considerable body of evidence has accumulated that speakers have detailed phonetic knowledge, not just categories of phonological theory (Bradlow 1995; Pierrehumbert 2000). Based on such evidence, the usage-based framework (Pierrehumbert 2001; Bybee 2000, 2001 among others) proposes that mental representations of phonological targets and patterns are gradually built up through experience with speech. This challenges standard models of phonology and phonetics at two levels. In all standard models, the lexicon is distinguished from the phonological grammar and thus the exact phonetic details are processed by the rules of grammar. As a result, the phonetic implementation component applies in exactly the same way to all surface phonological outputs. The purpose of the present paper is to show that phonetic implementation is variable and gradient, against the proposal the standard models put forth.

This paper focuses on the Korean speakers' processing of the English epenthetic stops between the sonorant and fricative segments (dense as [d?n*t*s], warmth as [w??m*p*?]) (Clements 1987; Ohala 1974; Fourakis & Port 1986; Warner & Weber 2001). As for the stimuli of the perception experiment, 120 English non-words with a mono-syllable structure of CVC_1C_2 were created, where C_1=/m, n, ?, l/, and C_2=/s, ?, ?/. The stimuli including additional 63 filler items with target stops as separate phonemes were recorded by two native English speakers and one representative data among them were given to two groups of native Korean speakers, those with high and low proficiency in English, as well as native English speakers. Using a phoneme monitoring task (Warner and Weber 2001; Connine and Titone 1997), I tested each group of subjects' processing: each group of subjects listened to speech and detect the target stops such as [p], [t], and [k]. The number of their responses were computed to determine how often subjects succeed in recovering the string of segments produced by the native English speaker.

The preliminary results of the experiment show that 1) English epenthetic stops were poorly identified by native Korean speakers with low English proficiency, even in the case where stimuli with strong acoustic cues were provided with, but 2) there's a high correlation between perception of epenthetic stops and speakers' English proficiency, showing the possibility of the improvement of processing abilities. The covert sound structures were shown to be poorly identified by foreign language learners, but their processing abilities were improved by learning. This study thus provides evidence that exact phonetic targets and patterns of variation must be learned during the course of language acquisition as argued by many proposals in the usage-based framework.

It was also shown that processing of phonetic variability is affected by several factors, including Korean phonotactic constraints prohibiting consonant clusters in the syllable coda, and the inherent acoustic properties of epenthetic stops, which led to the asymmetry in the perception of epenthetic stops, namely, perception of epenthetic [p] and [k] vs. [t] (Jun 1996; Hume et al. 1999).

Perception of cross-language vowel differences: A longitudinal study of native Spanish adults learning English

Satomi Imai*, James E. Flege* & Ratree Wayland†
University of Alabama at Birmingham and University of Florida†*

A great deal of recent research has focused on the ability of adults to learn to produce and perceive the phonetic segments of a second language (L2). The foreign accent that arises from errors in producing L2 vowels and consonants has often been attributed to inaccurate perception. A number of theorists have proposed that the perception of L2 phonetic segments will be inaccurate following attunement to the L1 phonetic system because L2 input is warped due to the use of L1-tuned filter weights, or because L2 phonetic dimensions that are non-contrastive in the L1 are filtered out (Trubetzkoy, 1939/1958; Sebastián-Gallés and Soto-Faraco, 1999; Kuhl, 2000; Iverson et al., 2001). According to the Speech Learning Model, on the other hand, adult learners of an L2 can in time accurately perceive the phonetic properties of L2 speech sounds (Flege, 1999, 2002, 2003). However, relatively little is known at present concerning adults' ability to detect cross-language phonetic differences and the time course of changes in the perception of L1-L2 phonetic differences.

This study evaluated the ability of 15 native Spanish (NS) adults to perceptually distinguish English and Spanish vowels. The aims of the study were to determine if the NS participants perceived a difference between English vowels and neighboring Spanish vowels, and whether perceived English-Spanish differences would increase as the NS participants gained experience in English

The NS participants were learning English as an L2 in the United States (U.S.). They had a mean age of 34 years, had arrived in the U.S. at an average age of 29 years, and had lived in the U.S. for an average of 3.8 years when tested the first time (Time 1). The perceptual stimuli consisted of 5 Spanish and 11 English vowels that had been spoken in a /b_b/ context by 5 native speakers of each language. The vowel stimuli were presented to the NS participants on 6 occasions, each separated by 1 year.

The Spanish and English vowel stimuli were randomly presented for classification as one of the 5 vowels of Spanish. Goodness of fit ratings were also obtained for each token (5 = good instance of a Spanish vowel, 1 = poor instance, 0 = not a Spanish vowel). The dependent variables were the percentage of times that the 5 tokens of each vowel category were classified as "not Spanish" and the goodness of fit ratings obtained for the remaining tokens.

The NS participants were unable to distinguish the English /i/ tokens from the Spanish /i/ tokens. The five English /i/ tokens were rarely identified as 'not Spanish" and received goodness ratings that were virtually indistinguishable from the ratings obtained for Spanish /i/, even at the end of the 5-year study. The NS adults also failed to distinguish the low central Spanish vowel /a/ from the two low vowels of English. However, the NS participants were able to distinguish the remaining English vowels of interest from neighboring Spanish vowels (i.e, the vowels in "bet" and "bait" from Spanish /e/; the vowel in "boat" from Spanish /o/; the vowel in "boot" from Spanish /u/).

The NS participants' ability to perceive Spanish-English vowel differences was related closely to the magnitude of acoustic differences between the five tokens of the various Spanish and English vowel categories. There was little change in the magnitude of perceived Spanish-English differences over the 5-year test interval. The one exception was English /u/. The NS

participants judged the English and Spanish /u/ tokens to be equally good instances of the Spanish /u/ category at Time 1, but gave significantly lower ratings to the fronted English /u/ than to Spanish /u/ one year later (Time 2) and in all subsequent sessions. Taken together, the results suggested that adult learners of an L2 perceive all but the smallest cross-language vowel differences, but that certain differences may not be noted immediately.

References

Flege, J. (1999). Age of learning and second-language speech. In David Birdsong (Ed.) New Perspectives on the Critical Period Hypothesis for Second Language Acquisition. Hillsdale, NJ: Lawrence Erlbaum. Pp. 101-132.

Flege, J. (2002). Interactions between the Native and Second-language Phonetic Systems. In P. Burmeister, T. Piske and A. Rohde (Eds.) An Integrated View of Language Development: Papers in Honor of Henning Wode. Trier: Wissenschaftlicher Verlag. Pp. 217-244.

Flege, J. (2003). Assessing constraints on second-language segmental production and perception. In A. Meyer and N. Schiller (Eds.) Phonetics and Phonology in Language Comprehension and Production, Differences and Similarities. Berlin: Mouton de Gruyter. Pp. 319-355.

Iverson, P., Kuhl, P. Yamada, R., Diesch, E., Tohkura, Y., Ketterman, A., and Siebert, C. (2001). A perceptual interference account of acquisition difficulties for non-native phonemes. Speech, Hearing and Language: Work in Progress, 13, 106-118.

Kuhl, P. (2000) A new view of language acquisition. Proceedings of the National Academy of Science, 97, 11850-11857.

Sebastián-Gallés, N. and Soto-Faraco, S. (1999) On-line processing of native and non-native phonemic contrasts in early bilinguals. Cognition, 72, 111-123.

Trubetzkoy, N. (1939/1958) Grundzüge der Phonologie. Göttingen: Vandenhoek & Ruprecht.

Articulatory correlates of sonority scale in consonants

Mafuyu Kitahara*, Tadao Murata† & Haruka Fukazawa†
Yamaguchi University and Kyushu Institute of Technology†*

Sonority sequencing has been considered as a guiding principle for syllabification process in phonology. The notion of sonority is loosely connected to "loudness", but its articulatory correlates have not yet been clearly determined. At a certain level of abstraction, articulatory "effort" might constitute an inverse scale of sonority. However, the degree of effort is not simply measurable in speech. Low-level kinematic parameters, such as the time course, the mass, and the velocity of articulators may differ in various ways for each consonant.

The present study examines whether there is an articulatory correlates of sonority by looking at a rather high-level variable: a switching point from VC to CV syllable when the VC syllable is repeated at an increasing rate. This rate induced resyllabification is originally noted by Stetson (1951). We look whether the point and the manner of articulation of consonants have any effect on the switching point from VC to CV.

Speakers were instructed to entrain to a metronome which increased tempo from 300 to 100 ms interval. [iC] and [aC] syllables were repeated where [C] includes [p,t,k,b,d,g,f,s,z,m,n,l]. About 20 speakers participated. Consonant and vowel onsets are measured with respect to metronome beeps.

Results show that the switch point measure averaged across natural classes closely corresponds to the sonority scale in (1).

(1) (a) voiceless stop < voiced stop
(b) stop < fricative
(c) obstruent < sonorant

That is, the switch point from [VC] to [CV] are earlier for the classes on the less sonorous side. This means less sonorous consonants more easily slip out of [VC] and fall on to [CV] syllabification. This interpretation is in accordance with the dynamical systems theory (Kelso, 1995) where [CV] syllables are more stable attractors than [VC] syllables in the state space.

References

Stetson, R. H. (1951) Motor Phonetics, 2nd Edition, North-Holland, Amsterdam.

Kelso, S. J. A. (1995) Dynamic Patterns: The Self-Organization of Brain and Behavior, MIT Press, Cambridge, MA.

Schwa-productions and hesitations by native French speakers during English L2 performance

Barbara Kühnert[*†] & Cécile Fougeron[*]
Institut du Monde Anglophone[†] and CNRS/UMR 7018 Université de la Sorbonne Nouvelle[*]

Broadly speaking, the neutral vowel schwa can be considered as a kind of homebase to which the tongue returns frequently in the course of speech. As such, it can be considered as an indicator of the overall articulatory setting of a language. In French, for example, the neutral vowel is said to be produced as a central vowel with some degree of lip-rounding, which however may often be fronted, while in English the vowel is said to be mid-central with the lips in a neutral position (Fougeron and Smith, 1999; Delattre 1965).

At the same time, the neutral vowel schwa is the most frequently occurring vowel in spoken English (Roach, 1991). Its correct pronunciation is therefore closely correlated with the perceived degree of authentic native-like speech. Given the difficulties that speakers of a second language (L2) encounter when producing L2 sounds that are 'similar' to sounds in their native language (i.e. sounds that are nominally said to be the 'same', but differ as to how they are realized on the phonetic surface; Flege 1997), inaccuracies in the phonetic realization of schwa might indeed be a major source of global foreign accent. In addition, given the relationship between the preferred articulatory setting and the often schwa-like nature of hesitations or 'filled pauses', it might not only be in schwa-productions but also in the phonetic quality of hesitations that the origin of L2 speakers is revealed.

The paper presents an acoustic analysis of English schwa-realisations produced by French speakers and compares them with their own schwa-productions in French and schwa-productions of native speakers of Standard Southern British English. The L2 subjects were (i) French undergraduate students in their third year of English studies, and (ii) French university instructors teaching English. In order to evaluate the influence of the speaking task on L2 performance, the speech material was collected in formal reading tasks as well as free conversations. The frequent hesitations that occurred in the casual task will also be examined, exploring the degree to which even very advanced L2 speakers switch back to their 'natural' articulatory setting when speaking a foreign language.

References

Delattre, P. (1965). Comparing the Phonetic Features of English, French, German and Spanish. Heidelberg: Julius Verlag.

Flege, J. (1997). English vowel productions by Dutch talkers. In J. Leather & A. James (eds.), Second-Language Speech. Berlin, New York: Mouton de Gruyter, 11-52.

Fougeron, C. & Smith, C. (1999). Illustrations of the IPA: French. Handbook of the International Phonetic Association. Cambridge: Cambridge University Press, 78-81.

Roach, P. (1991; 2nd edition). English Phonetics and Phonology. Cambridge: Cambridge University Press.

Metrical Computation in L2 Stress Acquisition: Evidence from Chinese-English Interlanguage

Shu-chen Ou & Mitsuhiko Ota
University of Edinburgh

Much research on the acquisition of second language (L2) stress by speakers of another stress system has converged on the view that the course of development is constrained by principles of metrical phonology (e.g., Archibald 1993, 1995; Pater 2000). However, it remains an empirical question whether this generalization is extendable to the acquisition of L2 stress by learners whose native language does not exhibit dynamic stress. In one of the few studies that investigate the L2 acquisition of a stress system by non-stress language speakers, Archibald (1997) concludes that Chinese and Japanese speakers do not compute metrical structures in learning L2 English, but instead depend on item-by-item lexical storage of the stress pattern. Contrary to this claim, the current study presents evidence that tone language speakers also engage in systematic computation of L2 stress based on the prosodic and morphosyntactic information of the target language.

Twenty Mandarin/Taiwanese speakers participated in a perceptual preference experiment with 24 English pseudo-words presented auditorily. The stimulus words had either two or three syllables with different syllable structures and were embedded in a carrier sentence either as a noun or a verb. The disyllabic words had stress on the initial or final syllable, and the trisyllabic words on the initial (antepenultimate) or penultimate syllable.

The results showed subjects' sensitivity to the syllable structure and its interaction with the morphosyntactic property of the word. In trisyllabic pseudo-nouns, penultimate stress was preferred when the syllable was closed by a consonant. The tendency was stronger when the post-vocalic consonant was a sonorant than an obstruent. This indicates that Chinese learners of English use syllable weight to determine the position of stress, and the decision is influenced by the sonority of the coda consonant, in accordance to the sonority-weight interaction attested in weight-sensitive stress systems. Furthermore, subjects preferred initial stress for CV.CVCC words when they were presented as a noun, but preferred final stress for phonologically identical words when they were presented as a verb. This preference was even stronger than that given by native-speakers of English who served as control subjects. Chinese speakers therefore are also aware of the systematic difference in the extrametrical materials between nouns and verbs in English.

The results do not tell us whether these patterns are extracted from the input in L2 or whether they reflect non-apparent transfer from the metrical computation that underlies Chinese tone phonology. One thing that is certain, however, is that the acquisition of L2 stress does follow the principles of metrical phonology even when the learner's L1 does not have a dynamic stress system.

The Production of Indonesian English: Does L1 Phonological Transfer Occur At All?

Ninik H. Poedjianto
University of Glasgow

It has been long recognized in studies of 'foreign accent' that one's pronunciation of the target language leaves traces of the phonology of his/her first language (L1) (Munro, 1998). To what extent the L1 phonology is transferred still requires further research. The circumstances become more complex when the L1 phonology has not been under scrutiny phonetically. This is certainly the case with the L1 in this paper, namely Indonesian. In terms of voicing contrast, for instance, Indonesian stops are characterized as voiceless and voiced phonologically (Moeliono and Dardjowidjojo, 1993); and these characteristics are also found in English. However, there is no agreement yet to what extent the vocal fold vibration occurs in Indonesian voiced stops (Alieva *et al.*, 1991).

The aim of this study is to investigate whether there is any transfer of L1 phonology in the production of English by Indonesian learners across different levels of proficiency and gender. This paper attempts to address the following questions:

1. What are the phonetic characteristics of voicing contrast in Indonesian English (IE) produced by Surabaya learners of English as a foreign language?
2. What are the phonetic characteristics of voicing contrast in Indonesian produced by the same subjects?
3. Is there L1 phonological transfer in the production of IE?

For the dataset, I recorded English and Indonesian speech from 30 students from 3 proficiency levels, 5 female and 5 male at each level in a private English language school in Surabaya, East Java, Indonesia. The recordings were made using a Sony TCD7 DAT recorder and a wide range frequency clip-on microphone, and were digitized at 16 kHz, 16 bits into xwaves. To answer the preceding questions, I focus on the aspect of voicing contrast of Indonesian English (IE) consonants as in /p/ and /b/. I examined 6 English monosyllabic words (i.e. 'path'-'bath', 'push'-'bush', and 'pass'-'bus') repeated twice each in a carrier phrase *Say __ again*, and 2 Indonesian words (i.e. 'pak'-'bak' (= *Sir-tub*)) repeated twice each in a carrier phrase *Kata __ ada* (= *There is __ word*). The analyses were conducted in terms of: narrow phonetic transcription, parametric representations from spectrographic observation, temporal measurements (i.e. VOT, duration of medial phase, and voicing during medial phase), spectral tilt measurements (i.e. H1-H2, H1*-H2*, H1-A3), and fundamental frequency (F_0) measurement.

Results from GLM repeated measures show voicing contrast in Indonesian (I) is significantly characterized by phonation (H1-A3). Subjects produced significantly different IE /p/ from I /p/, and IE /b/ from I /b/ in terms of VOT, voicing during medial phase, and phonation (H1-A3). For medial phase duration and F_0, subjects made a significant difference between IE /p/ - I /p/ and IE /b/ - I /b/ respectively. Despite their strategies to produce native English-like voicing contrast by distinguishing the realizations of the same phonemes (i.e. /p/ and /b/) in two languages (i.e. IE and I), the only two significant phonetic characteristics to contrast voicing in IE are VOT and F_0. In general, therefore, rather than transferring the L1 phonology to produce IE voicing contrast, Indonesian speakers of English learn new phonetic characteristics instead.

It is an interesting fact that Indonesian learners did not transfer the phonation feature in their L1 to contrast voicing of their English; and this will be discussed at the conference. In conclusion, this study supports Flege's hypotheses that different phonetic realizations of the same phonemes in L1 and target language prevent learners from producing native-like English

sounds (Flege, 1987), and that phonetic space is restructured during learning a foreign language (Flege, 1988).

References

Alieva, N. F., Arakin, V. D., Ogloblin, A. K. and Sirk, Y. H. 1991. Bahasa Indonesia: Deskripsi dan Teori. Yogyakarta: Kanisius. [Indonesian: Description and Theory]

Flege, J. E. 1987. The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification. Journal of Phonetics. 15: 47-65.

Flege, J. E. 1988. The production and perception of speech sounds in a foreign language. In Winitz, H. (ed.) Human Communication and Its Disorders: A Review. Vol. 3. Norwood, NJ: Ablex. pp. 224-401.

Moeliono, A. and Dardjowidjojo, S. 1993. Tata Bahasa Baku Bahasa Indonesia. Jakarta: Balai Pustaka. [The Standard Grammar of Indonesian]

Munro, M. J. 1998. The effects of noise on the intelligibility of foreign-accented speech. Studies in Second Language Acquisition. 20 (2): 139-154.

Infants' learning of unnatural rules

Amanda Seidl* & Eugene Buckley†
Purdue University and University of Pennsylvania†*

It has been often suggested that young infants prefer phonologically natural alternations over unnatural ones (McCarthy and Prince, 1995; Tesar and Smolensky, 1993, 1998, 2000;) and that there is ample evidence for this assertion in both production (Demuth, 1995; Gnanadesikan, 1995; Pater and Paradis, 1996) and perception (Jusczyk, Smolensky and Allocco, 2002). In the two experiments presented in this paper we exploited infants' sensitivity to the frequency distribution of speech sounds in the input (Maye, Werker and Gerken, 2002) in order to test whether infants really do learn phonologically natural rules or alternations more easily than unnatural ones, or whether naturalness plays no role in the learnability of rules by infants.

We used an established procedure for statistical grammar learning in infants (Chambers, Onishi, and Fisher, 2003; Gómez and Gerken, 1999; Marcus, Vijayan, Rao, and Vishton, 1999; Maye, et al., 2002; Saffran, Aslin and Newport, 1996). These studies, and ours, use the Headturn Preference Procedure to look at infants' use of transitional probabilities to identify grammatical units or groupings. In this method, infants are familiarized for approximately two minutes with a series of sentences or word lists that obey a certain grammatical rule. Then subjects are tested on sentences or wordlists that either obey or do not obey the familiarized rule. If learning of the rule occurs in these experiments we predict that the infants will show a novelty effect and attend longer to unfamiliar or novel wordlists in the test phase.

To avoid problems with the learning of these rules we tested English-learning infants at 8.5–9.5 months because this is an age at which adult-like rules and phonotactics are not yet firmly cemented (e.g., Werker an Tees, 1984). In addition, to avoid interference of native phonotactics, we use alternations that crucially do not occur in English.

In Experiment 1 we tested a manner alternation. Our natural rule was: Fricatives occur only intervocalically (as found in cases of intervocalic spirantization in many languages). Our unnatural rule was: Fricatives occur only word-initially (not attested in any language we know of). Half the infants were familiarized with the natural rule (Group 1) and the other half were familiarized with the unnatural rule (Group 2). Both groups were then tested on novel lists that were consistent with either the natural or unnatural rule. Results from 24 subjects revealed that while both groups of infants preferred to listen to word lists that were inconsistent with their group's familiarized rule ($p<.01$), there was no significant difference between the group learning the natural rule and the group learning the unnatural rule.

In order to find out whether this was a general property of phonological rule learning or just something specific to this type of rule, in Experiment 2 we tested a place alternation instead of a manner one. In this experiment, our natural rule was: Labial consonants are followed by round vowels, coronal consonants by front vowels (mimicking cases of consonant-vowel place assimilation). Our unnatural rule was: Labial consonants are followed by high vowels, coronal consonants by mid vowels (not attested). Once again we familiarized two groups of infants to the two different rule types. Preliminary results revealed that once again, while the infants were able to learn both rules, they showed no specific preference for the natural over the unnatural rule.

In sum, we found that even young infants were able to learn fairly subtle rules of manner and place features, and that infants performed equally well at rule learning in the natural and

unnatural conditions. We take this as evidence that it is possible for infants to learn unnatural phonological rules and evidence against active constraints against unnatural rules in the grammar of infants at this age.

The Role of Gestural Overlap in Perceptual Place Assimilation: Evidence from Korean

Minjung Son*‡, Alexei Kochetov*† & Marianne Pouplier$^{\Theta}$*
Yale University‡, Haskins Laboratories, Simon Fraser University† and University of Maryland$^{\Theta}$*

The question of the origins of place assimilation has been a subject of increasing interest and controversy [1, 2, 4, 7, 8, 9, 11]. While researchers disagree on specific factors that cause perceptual place assimilation, it is commonly regarded as an important source of phonological assimilation [1, 8]. In particular, there are opposing views on whether perceptual assimilation is exclusively attributable to gestural reduction [5, 6] or can be triggered by gestural overlap as well [1, 3, 12]. The present study investigates whether perceptual place assimilation is uniquely attributable to either gestural reduction or overlap.

As an argument for the gestural reduction hypothesis, Jun [5, 6] presented data on regressive place assimilation in Korean /pk/ clusters. Investigating changes of oral pressure during the production of /pk/, Jun found that some tokens displayed gradient gestural reduction but no overlap, while other tokens showed gestural overlap without any reduction. In a follow-up perceptual study with Korean and English listeners he found that partially reduced tokens of /pk/ were overwhelmingly perceived as assimilated [kk], while overlapped tokens did not undergo perceptual assimilation. However, Jun's methodology did not allow him to distinguish between a partially and a fully reduced lip gesture. His assumption that Korean assimilation is a gradient, postlexical process was based largely on theoretical grounds. Additionally, all the tokens he identified to be partially reduced contained the sequence /pkw/. The small changes in oral pressure observed during these tokens can thus not uniquely be attributed to a potentially reduced /p/ or coarticulation of the velar labial-glide sequence.

The present study reports articulatory movement data recorded with an EMMA system [10] during the production of /pk/ clusters by a single Seoul Korean speaker. Stimuli consisted of real words with /pk/ clusters word-medially and at word boundaries, produced at two speaking rates; nonwords with word-medial /pk/ clusters were also collected. The results show that in word-medial /pk/ clusters the lip gesture was either fully present (with varying degrees of overlap) or completely absent, regardless of speech rate. Notably /pk/ clusters across word boundaries and word-medially in nonwords exhibited substantial overlap, but no reduction of C1. Contrary to Jun's hypothesis, the results indicate that word-medial lip reduction during /pk/ in Korean is a categorical rather than a gradient, postlexical process (though its occurrence is optional, or stochastic). Crucially, Jun's listeners failed to identify a /p/ in reduced /pk/ clusters not because of perceptual place assimilation, but likely because it was not articulated by his speakers.

Using the production data obtained in the EMMA experiment, we conducted a phoneme identification experiment with Korean and English listeners, presenting them with VC and VCCV stimuli that exhibited either overlap or categorical reduction of the lip gesture. Data from 10 subjects show that listeners often failed to recover the lip gesture in overlapped tokens and perceived [kk]. This was observed in up to 39% of the cases for Korean listeners and in up to 36% for English listeners. Unsurprisingly, a fully reduced lip gesture in /pk/ was not recovered by listeners, but perceived as [kk]. While the results allow us to conclude that overlap does play a role in perceptual place assimilation independent of reduction, any unique contribution of reduction cannot be investigated on the basis of Korean. The results suggest that gestural overlap

can be regarded as an important factor in the evolution of language-particular patterns of assimilation [3, 8].

References

[1] Browman, Catherine and Louis Goldstein. 1990. Tiers in articulatory phonology, with some implications for casual speech. In: Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech, eds. J. Kingston and M. Beckman. Cambridge Univ. Press, 341-376.

[2] Byrd, Dani. 1992. Perception of assimilation in consonant clusters: a gestural model. Phonetica, 49, 1-24.

[3] Chen, Larissa Hoerning. (to appear) The origins in overlap of place assimilation. Proceedings of the 22nd West Coast Conference on Formal Linguistics (WCCFL 22), eds. M. Tsujimura and G. Garding. Somerville, MA: Cascadilla Press. 101-114.

[4] Hume, Elizabeth, and Keith Johnson. 2001. A model of the interplay of speech perception and phonology. In: The role of speech perception in phonology, eds. E. Hume and K. Johnson. New York, N.Y.: Academic Press. 3-26.

[5] Jun, Jongho. 1995. Perceptual and articulatory factors in place assimilation: an optimality theoretic approach. PhD Dissertation. UCLA.

[6] Jun, Jongho 1996. Place assimilation is not the result of gestural overlap: Evidence from Korean and English. In Phonology 13.3. 377-407.

[7] Jun, Jongho (to appear) Place assimilation. In: The Phonetic Bases of Markedness, eds. B. Hayes, R. Kirchner and D. Steriade. Cambridge University Press.

[8] Ohala, John. 1990. The phonetics and phonology of aspects of assimilation. In: Papers in Laboratory Phonology 7, eds. C. Gussenhoven, T. Rietveld, and N. Warner. Cambridge Univ. Press, 258-275.

[9] Nolan, Francis. The descriptive role of segments: Evidence from assimilation. In Paper in Labortory Phonology 2: Segement, gesture, and tone, eds. G. Docherty and G. D. Ladd. Cambrige: Cambrige University Press, 1991. 261-280.

[10] Perkell, J.S., Cohen, M., Svirsky, M., Matthies, M., Garabieta, I. and Jackson, M. 1992. Electro-magnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements. Journal of the Acoustical Society of America 92: 3078–3096.

[11] Steriade, Donca. 2001. Directional asymmetries in place assimilation: a perceptual account. The role of speech perception in phonology, eds. E. Hume and K. Johnson. San Diego: Academic Press. 219-250.

[12] Surprenant, Aimée and Louis Goldstein. 1998. The perception of speech gestures. Journal of the Acoustical Society of America 104: 518–529.

Testing Usage-based Phonology, Just for the /l/ of it.

Geoffrey S. Nathan
Wayne State University

In recent work Bybee (2000,2001) Pierrehumbert (2001,2002), and others have challenged the very center of phonological theory by arguing that phonemes are emergent generalizations from stored instances of individual pronunciations, and, in many cases surface forms may well not be computed on-line in the process of production and/or perception. In addition to cases of morphophonemic variation such as the variable deletion of 'past tense' /t,d/, Bybee suggests that even such prototypical instances of allophony as dark vs. light /l/ in English might not be connected in the minds of native speakers (Bybee 2001:87-88). In addition, she suggests that such storage issues are strongly affected by the relative frequency of the words being considered.

In this paper I will report on a acoustic study in which speakers' production is manipulated so as to examine whether they do in fact store allophones or compute them online. In a related part of the study I will study whether differences of word frequency lead to different results.

It is known that repetitions of words with canonical /VC/ form, produced with increasing speed, tend to be phonologically reanalyzed by speakers as /CV/, so that saying 'up up up' faster and faster leads to things more analagous to 'puh puh puh.' (Tuller and Kelso 1990) Speakers will be asked to repeat frequent words, such as 'all' and less frequent words such as 'L' in this manner. Spectrographic measurements of the /l/'s in these cases (as well as the vowel, which, for most English speakers has a different allophone before tautosyllabic /l/ than before other consonants) as they are reanalyzed as onset rather than coda will reveal (presumably) subconscious on-line processing if it actually does occur. Bybee/Pierrehumbert's theory predicts that there should be differences due to frequency effects. Furthermore, if taken at face value, there should be no effects on the /l/'s at all, regardless of rate, if the assumption is correct that speakers do not place onset and coda /l/'s in the same category.

References

Bybee, Joan. 2000. The phonology of the lexicon: Evidence from lexical diffusion. In M. Barlow & S. Kemmer (Eds.),Usage Based Models of Language (pp. 65-86). Stanford: Center for the Study of Language and Information.

Bybee, Joan. 2001. Phonology and Language Use. New York, Cambridge: Cambridge University Press

Pierrehumbert, Janet. 2001. Exemplar dynamics: Word frequency, lenition and contrast. in Frequency and the emergence of linguistic structure. ed. by Joan Bybee, Paul Hopper. Amsterdam: John Benjamins.

Pierrehumbert, Janet. 2002. Word-specific phonetics. in C. Gussenhoven and N. Warner eds. Laboratory Phonology VII, Mouton de Gruyter, Berlin. 101-140

Tuller, B., & Kelso, J. A. S. (1990). Phase Transitions in Speech Production and Their Perceptual Consequences. In Jeannerod (Ed.),Attention & Performance XIII , (pp. 429-451).

Contrast enhancement in Croatian clear speech

Rajka Smiljanic & Ann Bradlow
Northwestern University

Under adverse listening conditions, talkers naturally and spontaneously adopt a distinct, intelligibility-enhancing mode of speech production, known as "clear speech." Previous work on English clear speech production and perception has established that (1) naturally produced clear speech enhances intelligibility for various listener populations under degraded listening conditions and (2) speakers adopt a wide range of acoustic/articulatory adjustments in order to facilitate listeners. These conversational-to-clear speech modifications include a slower speaking rate, a wider dynamic pitch range, greater sound pressure levels, more salient stop consonant releases, expansion of vowel space, etc. An open question is how much these modifications reflect language-general, signal enhancement strategies and how much they reflect language-specific, structural enhancement strategies. Cross-language comparative data are crucial for addressing this question. More generally, such comparisons provide a window into the interaction of lower-level acoustic-auditory and higher-level phonological factors in spoken language processing.

Accordingly, the goal of the present study was to examine clear speech production and perception in a language whose sound structure differs substantially from English, namely Croatian with a phonemic vowel length contrast. Two native speakers of Croatian, one male and one female, read 20 sentences in conversational and clear speech modes. A sentence-in-noise perception test was conducted in order to investigate whether the articulatory adjustments implemented by these two speakers will confer an intelligibility benefit for native Croatian listeners. The results showed a clear speech intelligibility effect for all listeners. Similar to the benefit of clear speech found in English, the Croatian listeners were better at identifying words produced in the clear speech mode than in the conversational speech mode. Additionally, a talker effect was found: The clear speech intelligibility effect was greater for the male talker than for the female talker.

Acoustic analyses showed that for both speakers a decreased speaking rate, inclusion of more frequent pauses and a wider pitch range characterized the clear speech mode relative to the conversational speech mode thereby enhancing the overall acoustic salience of the signal. Other adjustments, such as, vowel space expansion and longer stop closure durations, characterize the clear speech productions as well. Crucially, however, the two speakers implemented different strategies in dealing with the lexical vowel length contrast. The male speaker lengthened long and short vowels in an asymmetrical way (short vowels were lengthened by 53% and long vowels by 75% in clear speech). In this way, the length contrast was exaggerated in clear speech. In contrast, the female speaker lengthened both long and short vowels to a similar degree (short vowels were lengthened by 51% and long vowels by 41%). For this speaker, the amount of the phonemic length contrast remained similar in both speaking styles. These results showed that both speakers enhanced the overall acoustic salience of the signal (slower, wider pitch range) but only one speaker enhanced the language specific phonemic vowel length contrast. It is argued that such language specific, structural enhancement contributes to the observed asymmetry in the intelligibility benefit between the male and female talkers.

The results of the listening test suggest that enhancing the signal as well as the language-specific phonemic contrast yields a greater intelligibility benefit than just enhancing the signal.

Perception and production of clear speech seems to be guided by principles that reflect both lower-level acoustic-auditory and higher-level phonological factors in spoken language processing.

The development of perceptual cue weighting within and across monosyllabic words

Catherine Mayo & Alice Turk
University of Edinburgh

Numerous studies have shown that adults and children weight acoustic cues differently in perception of certain speech contrasts. For example, in identifying /su/ versus /shu/, children give relatively more weight to within-syllable vowel formant transitions than do adults (Nittrouer & Studdert-Kennedy, 1987). In contrast, children have been found to give less weight than adults to across-syllable vowel formant transitions in identifying the contrasts /arda/ versus /arga/ and /alda/ versus /alga/ (Nittrouer, 1992). Nittrouer and colleagues have proposed that these adult-child cue weighting differences are the result of perceptual strategy differences between adults and children. According to this theory, children process speech in terms of relatively large units, such as syllables or monosyllabic words, while adults process in terms of smaller units. Nittrouer and colleagues claim that listeners give more weight to within-processing unit cues, and less weight to cues that cross processing unit boundaries (Nittrouer & Miller, 1997). Therefore, as found in the studies described above, children should be biased towards syllable-internal transitions, while giving less weight to syllable-external, or across-syllable transitions.

However, recent studies (Mayo & Turk, submitted) have found that adult-child differences in cue weighting are affected by segmental context. It is possible, therefore, that the different patterns of cue weighting observed by Nittrouer for within- and across-syllable transitional cues could be due to the segmental differences between the contrasts examined (fricative-vowel compared to approximant-stop). The aim of the current study was to address this issue by examining five-year-olds' and adults' weighting of phonetically identical vowel-formant offset transitions (from the vowel /e/ into a following /b/ or /d/ closure) in both within- and across-monosyllabic-word contexts: "Abe E" versus "Ade E" and "A bee" versus "A dee". Listeners were explicitly told the location of the word boundary in each case. Results from five-year-olds indicate that children's patterns of weighting are not as consistent as previously observed. While some children did give more weight to within-monosyllable transitions and less weight to across-monosyllable transitions, an equal number of children gave more weight to across-monosyllable transitions and less weight to within-monosyllable transitions. This would suggest that children are sensitive to the placement of acoustic cues within or across word/syllable boundaries. The findings are not consistent with Nittrouer's view that children always weight within-syllable cues more heavily than across-syllable cues. These results are, however, consistent with the view that 5-year-olds are in the process of learning that cues behave differently within and across word/syllable boundaries. Results from adult listeners will be used to test this hypothesis.

References

Mayo, C. & Turk, A., (submitted), `Adult-child differences in acoustic cue weighting are influenced by segmental context: Children are not always perceptually biased towards transitions'.

Nittrouer, S. (1992), `Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries', Journal of Phonetics 20, 351-382.

Nittrouer, S. & Miller, M. E. (1997), `Predicting developmental shifts in perceptual weighting schemes', Journal of the Acoustical Society of America 101, 2253-2266.

Nittrouer, S. & Studdert-Kennedy, M. (1987), `The role of coarticulatory effects in the perception of fricatives by children and adults', Journal of Speech and Hearing Research 30, 319-329.

Local acoustic cues distinguish two levels of prosodic phrasing: Speech corpus evidence

Tae-Jin Yoon, Heejin Kim & Sandra Chavarría
University of Illinois at Urbana-Champaign

In Pierrehumbert's model of American English intonation (1980), the sequence of phrase accent and boundary tone is used as a complex tonological mark of phrasal juncture. A modification is introduced by Beckman and Pierrehumbert (1986) who propose that "the phrase-accent plus boundary tone configuration in Pierrehumbert (1980) should be reanalyzed as involving correlates of two levels of phrasing," and they introduce the intermediate phrase as a level below the intonation phrase in the prosodic hierarchy. This analysis of an independent intermediate phrase is based on the auditorially perceived degree of disjuncture and on observed F0 contours from a small set of data, including controlled speech data produced under laboratory conditions and synthesized speech data. The analysis of intonation in speech produced outside of the experimental setting requires a transcription of the tone events that mark prominence and phrase juncture, and much of the recent research relies on the intonation transcription standard of the ToBI (Tones and Break Indices) system for prosody labeling (Beckman & Ayers1997). The ToBI system adopts the Pierrehumbert-Beckman model, including the distinction between the intermediate and intonational phrase levels. However, in actual labeling practice with non-laboratory speech, it can be very difficult to judge the level of phrase juncture in cases where the pitch contour alone is not decisive. For example, a phrase ending in a low or falling pitch contour could be analyzed with a low intermediate tone (L-) or with a sequence of a low intermediate tone followed by a low intonation phrase tone (L-L%), as noted in the ToBI labeling guidelines. In such cases the human labeler must rely on cues other than the gross pitch contour to identify the level of phrasal juncture. This labeling challenge can be especially acute in conversational speech styles for speakers who exhibit less overall pitch variation than is often found in laboratory speech.

Our current research addresses two questions: Do human labelers reliably distinguish two levels of phrase juncture in non-laboratory speech? If so, what are the acoustic factors that condition the perceived distinction? We analyze intonation in two large speech corpora: the Boston University Radio News corpus of read speech, and the Switchboard corpus of telephone conversation speech. ToBI labels are available for the radio news corpus, and we have produced our own ToBI labels for a subset of Switchboard. Published reliability studies (including Yoon et al. (ms)) show that human labelers do reliably distinguish the intermediate from intonational levels of phrase juncture. In this paper we present our findings on the acoustic correlates of phrase juncture through a comparison of low-toned intermediate and intonational phrase boundaries (L- and L-L%) in Radio News and Switchboard speech. Normalized and raw acoustic measures of F0, intensity, and duration were taken from the phrase-final syllable rime for each boundary type. The results show that in both corpora there are significant acoustic correlates of phrase level expressed in the phrase-final syllable rime. In both corpora, nucleus duration is longer (corroborating the finding of Wightman et al. 1992), and the F0 value at rime end is lower at intonation phrase boundaries compared to intermediate phrase boundaries. The read speech further exhibits significant differences in F0 drop and F0 slope over the final rime. Differences in intensity are observed in measures of peak intensity for both corpora, and somewhat more variably, in rime-end intensity. Our findings provide important empirical support from non-laboratory speech for the Pierrehumbert -Beckman model in its distinction of two levels of

phrase juncture. Our finding of acoustic correlates of phrase level in the phrase-final rime, and most often at the rime-end, offers critical support for the claim that prosodic features are locally rather than globally associated in phonological structure (Beckman and Ayers 1997). Our F0 analyses based on local pitch normalization (after Patterson 2000) offer further specific confirmation for the locality of pitch cues to prosodic structure.

References

Beckman, M. and G. Ayers. 1997. *Guidelines for ToBI labeling* (version 3.0). ms. The Ohio State University.

Beckman, M. and J. Pierrehumbert. 1986. Intonational structure in Japanese and English. *Phonology Yearbook* 3, pp. 255-309.

Patterson, D. 2000. *A linguistic approach to pitch range modeling*. PhD dissertation, University of Edinburgh.

Pierrehumbert, J. 1980. *The phonetics and Phonology of English Intonation*. PhD dissertation, MIT.

Wightman, C., S. Shattuck-Hufnagel, M. Ostendorf, and P. Price. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America*, 91(3), 1707-1717.

Yoon, T., S. Chavarría, J. Cole, and M. Hasegawa-Johnson. Intertranscriber Reliability of Prosodic Labeling on Telephone Conversation Using ToBI. (ms.) [http://prosody.beckman.uiuc.edu/pubs/Yoon-etal-ICSLP2004.pdf]

Acoustic invariant approach and speaker-independent features of speech sounds (Ukrainian and English)

Maksym Vakulenko
Institute for Theoretical Physics-Ukrajina and The University of Texas-Arlington

A great advance has been made in acoustic phonetics in the second half of the 20th century. Particularly, the relevant monograph by Kenneth Stevens [1] clearly reflects the modern state of the art. There are many automatic speech recognition systems present on the market, being able to satisfy more or less whimsical requirements. However, the problems of sex, noise, learning, emotional state, etc. are not solved yet completely (see [2-4]). For comparison: a dog can easily distinguish its name in the noisy street whoever has pronounced it, aloud or in whisper, and perceives intonation, and is much quicker in learning.

The most remarkable feature of this research is application of the Acoustic Invariant Speech Analysis (AISA) methodology.

The vowels were pronounced in a normal tone, in whisper and in a changing tone (singing). This procedure allows one to filter out the characteristic independent on the way of pronunciation. Thus each Ukrajinian vowel - [a], [o], [u], [i], [y], [e] - turned out to have individual formant ratio across different speakers (r = 3/2 for [o], for example).

When the tone pitch is increased, the corresponding formant pairs move higher in the frequency scale. For example, we observed the main pair of persistent (invariant) formants Fp1 = 3000 Hz (2800–3200 Hz), Fp2 = 4000 Hz (3800–4200 Hz) and the additional one Fp1(2) = 6000 Hz (5700–6300 Hz), Fp2(2) = 8000 Hz (7700–8300 Hz) corresponding to the main tone frequency F0 = 800 Hz.

The formants of various acoustic realization of the English phoneme /i/ ([i:] and [I]) were found to have relatively stable persistent frequencies: Fp1 = 2000 Hz (1800 – 2200 Hz), Fp2 = 2500 Hz (2300 – 2700 Hz), with a formant ratio r = Fp2/Fp1 = 5/4 (large tertion) that corresponds to the Ukrajinian [i].

The separate sound [x] appears to have main resonance frequency 1000-1600 Hz (1300 Hz) and a secondary one 4000-4600 Hz (4300 Hz) with a smaller amplitude. The voice component is absent, as it should. However, the situation is drastically changed when [x] is followed by the vowel. It was noticed in [5-7] that in the marginal region between [x] and the nextcoming vowel, there arise an “aftersound” that has the voice component and its own formants. The afterglide that always appears before vowels and "open" pauses, manifests the syllable openness. It is analogous to the "dumb e" in French or to "er" ("hard sign") and "erchik" ("soft sign") in the ancient Slavic language. Let us emphasize that since afterglide is not necessarily conditioned by the presence of the nextcoming vowel, this phenomenon is relevant to the given consonant and should be distinguished from the onset of the following sound.

This effect was experimentally observed in each combination "consonant + vowel". Such “afterglide” between [x] and [a] has its own spectrum, with the leading overtone frequencies 720 Hz and 600 Hz. So it is different from [a] and should be regarded therefore as an interaction (coarticulation) phenomenon.

The coarticulation effects inherent to the speech flow, result in the formants variations in the transition region (naturally, the resulting sound characteristics are changed, too). Being

important for the speech perception, they require special investigation and will be discussed in future.

This work was partially supported by the Fulbright Program grant # 68427736.

References

1. Stevens K. N. Acoustic Phonetics. – MIT Press, 1998. – 607 p.
2. Flanagan James L. Speech analysis, synthesis and perception. Springer-Verlag Berlin – New York, 1965.
3. Ladefoged Peter. A Course in Phonetics. University of California, 1975. – 196 p.
4. Trends in Speech Recognition / Wayne A. Lea, Editor. – Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1980.
5. Vakulenko Maksym, Vakulenko Olegh. Ukrajinsjkyj pravopys: poghljad z Ukrajiny (Ukrainian spelling: view from Ukraine) // Naukovi zapysky Akademiji nauk vyshhoji shkoly Ukrajiny (Scientific reports of the Higher School Academy of Sciences of Ukraine). - Vol.4. - Kyjiv-Khreshhatyk, 2002. - P.129-138.
6. Vakulenko M. O. Akustychni kharakterystyky ta invarianty zvukiv ukrajinsjkoji movy (Acoustic characteristics and invariants of the Ukrajinian sounds). – Naukovyj visnyk kafedry JUNESKO KDLU (Scholarly News of the KSLU UNESCO Chair). – Filologhija, pedaghoghika, psykhologhija (Philology, Pedagogics, Psychology). – Vol. 1. Kyjiv, 2000. – P. 62–66.
7. Vakulenko M. O. Analiz i syntez zvukovykh spektriv ljudsjkogho movlennja / Analysis and synthesis of the sound spectra of human speech. – Puljsar, # 6-7, 1999. – P. 20-23. (in Ukrainian).

Why Some Things are Um-bearable: Investigating Perceptual and Articulatory Influences on Place Assimilation Processes

Steve Winters
Indiana University

Several cross-linguistic surveys have revealed that place assimilation processes target nasals more often than stops in the languages of the world. These surveys have also revealed that there is an implicational relationship between nasals and stops as targets of place assimilation: stops never undergo this process in any given language unless nasals do so, as well. Some researchers (e.g., Jun 1995, Boersma 1998) have hypothesized that there are perceptual reasons behind the asymmetric behavior of nasals and stops in place assimilation, suggesting that nasals are more susceptible to this process because they have weaker acoustic cues to their place of articulation.

This hypothesis was tested by investigating the ability of Dutch and English listeners to discriminate between nasals and stops of varying places of articulation in an AX discrimination experiment. Speakers of these two languages were chosen because Dutch and English exhibit different assimilatory patterns with respect to nasals and stops in connected speech. In Dutch, only coronal nasals are susceptible to place assimilation (Booij 1995), while in English, both coronal nasals and stops may undergo this process (Barry 1991). The listeners in this experiment were asked to discriminate between consecutive VCCV sequences; the consonants in these sequences were either nasal-stop or stop-stop clusters. These nasals and stops were all produced at either the labial, coronal, or dorsal points of articulation. The analysis of listeners' ability to discriminate between these VCCV sequences focused on VCCV pairs which differed in the place of articulation of the initial consonant, since these consonants are the most likely targets of place assimilation in English and Dutch. Listeners more easily discriminated stops than nasals in this context when the stops had audible release bursts; without these cues, stops did not have a perceptual advantage over nasals. The language of the listener had no significant effect on the discriminability of place in either nasals or stops. These results suggest that the hypothesized perceptual advantage of stops over nasals is mediated by the likelihood of speakers to produce them with audible release burst cues.

A subsequent experiment investigated the possibility that the difficulty of articulating consonant clusters consistently and accurately might account for nasals' comparative susceptibility to place assimilation. To test this hypothesis, Dutch and English speakers participated in a production experiment wherein they listened to and attempted to imitate the VCCV stimuli that were used in the AX discrimination experiment. The results of this study showed that speakers of both languages exhibited less accuracy in reproducing nasals than stops. Examining the acoustic variability of subjects' spoken repetitions also revealed an interaction between manner of articulation and the language of the speaker: nasal-stop sequences had greater durational variability than stop-stop sequences for only Dutch speakers—for English speakers, the durational variability of these sequences was the same. This articulatory variability seems to correspond to the different assimilatory patterns in the two languages. Together, these results suggest that the cross-linguistic asymmetry between nasals and stops as targets of place assimilation may be motivated more by the difficulty of articulating nasals in consonant clusters than by listeners' relative inability to perceive their place of articulation correctly.

References Cited

Barry, M. 1991. Temporal modelling of gestures in articulatory assimilation. Proceedings of the XIIth International Congress of Phonetic Sciences **4**, 14-17.

Boersma, P. 1998. Functional phonology: formalizing the interactions between articulatory and perceptual drives. PhD Dissertation, University of Amsterdam.

Booij, G. 1995. The Phonology of Dutch. Oxford: Oxford University Press.

Jun, J. 1995. Perceptual and articulatory factors in place assimilation: an optimality theoretic approach. PhD Dissertation, UCLA.

Subphonemic tonal distinction in Cantonese

Alan C. L. Yu
University of Chicago

Introduction: Past research has indicated that under some conditions, speakers produce small, but significant, phonetic differences in words which, based on processes of positional neutralization in phonology of the language, would be expected to be homophones (e.g., final voicing neutralization (Dinnesen & Luce 1984, Port & O'Dell 1986, Port & Crawford 1989, Tieszen 1997, Warner et al. In press, Yu Forthcoming); tonal neutralization (Peng 2000)). Some authors suggest that these kinds of subphonemic differences are due to orthographic differences or speaking style (Fourakis and Iverson 1984, Jassem and Richter 1989). ***This paper presents evidence of subphonemic tonal differences in Cantonese.*** Cantonese is a prime candidate for investigating the nature of subphonemic differences in speech production and perception since speakers do not possess conscious knowledge of the tonal phonology of the language. Cantonese has no standard orthography. While Standard Chinese orthography, which does not indication tone or durational differences, is part of the education curriculum, the romanization (i.e. Pinyin) is not taught in school. The results of this study have serious implications for theories of phonological representation.

Background: Cantonese has six tone classes (i.e. 55, 11, 35, 13, 33, 22). Besides the lexical 35 tone, mid-rising tones may also come about as the result of a deverbal nominalization process that derives mid-rising tones from semantically related level tone syllables (e.g., tsʰatˀ33 'to brush' → tsʰatˀ35 'a brush') and a sandhi process that reassociates a high tone of an optionally deleted syllable to the preceding level tone syllable (e.g., **tsʰatˀ33** jʌtˀ55 tsʰatˀ33 → **tsʰatˀ35** tsʰatˀ33). The following experiment was performed to investigate the production of these three types of mid-rising tones.

Methodology: Two subjects were asked to recite a list of Cantonese target disyllabic words/phrases in the carrier phrase /ŋɔ tɔkˀ ___ pɛi nɛi tʰæŋ/ 'I say __ for you'. The target words were ten quadruplets of CVV Cantonese syllables. The quadruplet consisted of segmentally identical CVV syllables with four different tonal types (i.e. M(orphologically-derived)-35, S(andhi-derived)-35 and L(exical)-33 & 35)

Results: The results indicate that the high targets of the three 35 tones are significantly different from each other, despite the fact that, according to the consensus view of Cantonese phonology (e.g., Chen 2000, Yip 2002), their phonological representations are identical, namely, sequences of MH level tones. The high target of a S-35 is higher than that of the M-35, which in turn is higher than the higher target of a L-35.

Conclusion: The results of this study demonstrate that phonological representation must include fine grain phonetic details (Pierrehumbert 2003) in order to capture these subtle, but systematic, phonetics differences in speech. We argue here that the subphonemic tonal difference observed between the L-35 and S-35 can be captured representationally by assuming that L-35 forms a unit (Yip 1989), while S-35 is composed of a sequence of MH level tones as predicted by a delinking and reassociation analysis of tone sandhi. The tone sequence approach to S-35 is supported by the fact that the pitch value of the high target of a S-35 does not significantly differ from that of a lexical 55. The distinct high target of a M-35 can be explained by the fact that the deverbal noun construction is a marginally productive synchronic process. The current phonetic

realization of M-35 might be a reflex of an earlier stage when a subset of deverbal nouns have L-35 representation, while others are productively derived in a way similar to the S-35.

References

Chen, Matthew (2000) Tone sandhi. Cambridge University Press.

Dinnsen, Daniel A. and Charles-Luce, Jan (1984) "Phonological neutralization, phonetic implementation, and individual differences," Journal of Phonetics 12.49-60.

Fourakis, Marios and Iverson, Gregory K. (1984) "On the 'incomplete neutralization' of German final obstruents," Phonetica 41.140-149.

Jassem, Wiktor and Richter, Lutoslawa (1989) "Neutralization of voicing in Polish obsturents," Journal of Phonetics 17.317-325.

Peng, Shu-hui (2000) "Lexical versus 'phonological' representations of Mandarin Sandhi tones," In Michael B. Broe and Janet B. Pierrehumbert (eds.) Papers in Laboratory Phonology V: Acquisition and the Lexicon. Cambridge: Cambridge University Press.

Pierrehumbert, Janet (2002) "Word-specific phonetics," In C. Gussenhoven, and N. Warner (eds.), Laboratory Phonology VII, Mouton de Gruyter, Berlin. 101-140.

Port, Robert F. and O'Dell, Michael L. (1985) "Neutralization of syllable-final voicing in German," Journal of Phonetics 13.455-471.

Port, Robert and Crawford, Penny (1989) "Incomplete neutralization and pragmatics in German," Journal of Phonetics 17.4.257-282.

Tieszen, Bozena (1997) Final stop devoicing in Polish: an acoustic and historical account for incomplete neutralization. PhD dissertation. University of Wisconsin, Madison.

Warner, Natasha, Jongman, Allard, Sereno, Joan, and Kemps, Rachèl (In press) "Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch," Journal of Phonetics.

Yip, Moira (1989) "Contour tones," Phonology 6.1.149-174.

Yip, Moira (2000) Tone. Cambridge: Cambridge University Press.

Yu, Alan C. L. (Forthcoming) "Explaining final obstruent voicing in Lezgian: Phonetics and history," Language.

Phonetic motifs and their role in the evolution of sound structure

Elinor Payne
University of Cambridge

The picture emerging from a growing body of research (e.g. Carter, 2003, Nolan, 1999, Barry 1992) is that there is a high degree of language-specific variability in phonetic interpretation. This variability raises questions about the internal architecture of sound systems and the dividing line between 'linguistic' and 'performance' knowledge. Being neither predictable nor random, the facts of a language's phonetic interpretation are structural properties of that language and therefore must form part of its users' linguistic knowledge, yet this remains distinct from what is supposedly a 'mentalist' knowledge of phonological contrast. Intuition persuades us that the dichotomy of 'sound events in the physical world' and 'sound events in the mind' is incontestable, but other notions, such as structure, form and meaning and how these map on to this dichotomy are more elusive. In this paper I propose modelling linguistic-phonetic detail as phonetic *motifs* and discuss how these motifs can be a source of structural innovation which may become ingrained in higher-level sound structure. I illustrate my discussion with a comparison of fine phonetic detail observed in consonant clusters and consonant gemination in different regional varieties of Italian.

Unlike constraints generating phonological patterns (as in e.g. Optimality Theory), motifs are descriptions of recurrent language-specific *low-level phonetic* events (though they may interact to give the *percept* of more abstract constructs, e.g. a syllable). The events described by motifs may be relatively localised (for example specifying sub-phonemic differences in place, manner, timing and phonatory setting) or occur over a longer domain, as with, for example, co-articulation strategies, articulatory settings and indices of phonetic prominence. As norms of behaviour which *distinguish* particular speech communities, they arguably constitute cultural knowledge. Culture is not, however, random: if language A executes a sequence of consonants with a greater degree of coarticulatory accommodation than language B, this is a non-arbitrary property of language A. Certain motifs may become associated with particular speech styles and/or boundaries, giving rise e.g. to connected speech processes. Motifs need not be discretely defined, but may vary around an average, allowing for context- and speaker-dependent variation.

In the second part of my paper I demonstrate how motifs may be a source of structural innovation. A given motif or combination of motifs may yield, epiphenomenally, more complex structures such as epenthesis, devoicing, assimilation, lengthening, secondary articulations, etc. While these 'structures' remain functionless they are sound 'junk' (after Lass, 1990), but they may come to be exploited for new, unplanned purpose (cf Lass's notion of linguistic exaptation developed in the context of morphosyntax). For example, gestural lengthening (through phonetic prominence) or coarticulatory accommodation of clusters may yield long consonants, which remain junk unless and until exapted for a specific function e.g. lexically contrastive geminates, or become more structurally ingrained, as in postlexical syntactic doubling. Structure may also act conservatively to maintain its 'shape', affecting the likelihood of exaptation (cf Kiparsky's 1995 notion of structure dependency). This source of phonological change is distinctly (phonetic-)linguistic in that innovation comes from within the (phonetic-)linguistic structure and not, directly, from the physics (e.g. Ohala, 1983, 1993) or functions of speech (e.g. Lindblom, 1986, 1990, Lindblom and Maddieson, 1988). It is also non-universal, in that innovation is unique to that language, and not reliant on any notion of universals, be they phonological or

phonetic. However, linguistic-phonetic shaping of sound systems is not incompatible with physical and functional accounts, and indeed in a more complete picture the three would interact.

References

BARRY, M. (1992) "Palatalisation, assimilation and gestural weakening in connected speech" in Speech Communication 11, 393-400

CARTER, P. (2003, forthcoming) "Extrinsic phonetic interpretation: spectral variation in English liquids" in J. Local, R. Ogden and R. Temple (eds), Papers in Laboratory Phonology VI, Cambridge, CUP

KIPARSKY, P. (1995) "The phonological basis of sound change" in J. Goldsmith (ed.) The Handbook of Phonological Theory, pp. 640-70, Oxford, Blackwell

LASS, R. (1990) "How to do things with junk: exaptation in language evolution" in Journal of Linguistics, 26 pp. 79-102

LINDBLOM, B. 1986 "Phonetic universals in vowel systems" in J. J. Ohala and J. Jaegar (eds) Experimental Phonology, Orlando: Academic Press, 13-44

LINDBLOM, B. 1990 "Explaining phonetic variation: a sketch of the H & H theory" in W. J. Hardcastle and A. Marchal (eds) Speech production and speech modelling, Dordrecht: Kluwer Academic Publishers, 403-40

LINDBLOM, B. and I. Maddieson 1988 "Phonetic universals in consonant systems" in L. M. Hyman and C. N. L. Routledge (eds) Language, speech and mind: studies in honor of Victoria. A. Fromkin, London: Routledge, 62-80

NOLAN, F. (1999) "The Devil is in the Detail", Proceedings of the 1999 ICPhS, San Francisco, Vol. 1, 1-8

OHALA, J. J. 1983 "The origin of sound patterns on the cognitive representation of speech" in T. Myers, J. Laver and J. Anderson (eds) The Cognitive Representation of Speech, Amsterdam: North-Holland, 111-22

OHALA, J. J. 1993 "Coarticulation and Phonology" in Language and Speech, 36, 155-70

Processing English [s] into Korean: Prosodic influence on sound change

Hyunsook Kang* & Seok-Keun Kang†
Hanyang University and Wonkwang University†*

English onset /s/ is matched with two different alveolar fricatives in Korean, plain [s] and tense [S], when perceived. That is, if English /s/ appears before a vowel in an English word, it is perceived as tense [S] in Korean, whereas if English /s/ appears before other consonant(s) in the onset cluster, it is perceived as plain [s] in Korean. Thus, Korean speakers perceive [Seil] for English 'sale' whereas they hear [sImail] for the English word 'smile.' ([I] is barred-[i].)

We show that this sound change is due to the combination of the following facts: Korean tense [S] is followed by a high-pitched vowel whereas Korean plain [s] is followed by a low-pitched vowel, the initial portion of which is voiceless due to the aspiration of Korean [s] and thus, when Koreans hear English [s], which is not perfectly matched with either Korean tense [S] or plain [s], they delay their judgment of the segment until they hear the prosodic characteristics of the following segment. Thus, when English /s/ is positioned in the context where the following segment is a high-pitched stressed vowel, it is identified as Korean tense [S]. If, however, English /s/ appears before another consonant, and thus, the consonant [s] itself should be perceived as a syllable due to the Korean syllable structure condition, Korean speakers perceive [sI], plain [s] with a voiceless vowel [I], for English [s].

As supporting evidence for our arguments, we first present that /s/ in 'sV-' and /s/ in 'sCV-' in English are not different from each other in the perception of Korean speakers. We cut the /s/s from these two types of words and got a perception test (12 native Koreans). The results show that regardless of the origin of English [s]s, Korean speakers did not show any significant tendency in perceiving them either as Korean tense [S]s or plain [s]s. Next, we performed a phonemic gating experiment, which consisted of several gates that contained [s] and varying portions of the following segment. The results show that Korean subjects did not show any significant tendency in perceiving short gates, containing consonant [s] and a very small portion of the following consonant in the onset cluster, to be either plain [s] or tense [S] just like their judgements for the isolated English [s]s. However, when more information of the second consonant became available in larger gates, a perceptual shift occurred and English [s] was increasingly judged to be Korean plain [sI]. If the gates contained English [s] and some portion of the following stressed vowel, however, more than 90% of the subjects replied that [s]s were perceived as Korean tense [S] as expected. Interestingly, when gates containing both the [s] and a portion of the stressless vowel with vocalic pulses were heard, 70% of Korean subjects responded that they perceived Korean tense [S]. We suspect that if English [s] is followed by a high-pitched stressed vowel, it is perceived as tense [S] more than 90% since the high-pitched stressed vowel in English has similarities to the high-pitched vowel after Korean tense [S]. We also think that Koreans perceive plain [s] from English [s] only if it is followed by enough of a portion of the following consonant to make the preceding English [s] be perceived as a syllable with a voiceless vowel followed by a more sonorous or high-pitched burst consonant, which would make the preceding [s] perceptually less salient. If English [s] is followed by a stressless vowel with vocalic pulses, more subjects perceived it as tense [S] but its proportion is much lower than that for English [s] followed by a stressed vowel due to the lack of the high pitch.

We also performed another perception test with the synthesized tokens in which /s/s from 'smile' and 'sale' were put in front of 'V' and 'CV' types of words, respectively. Again, regardless of the origin of English /s/s, the position of /s/ in the word decides its perception.

In conclusion, our findings support the argument that phoneme identification is delayed until other sounds in the surrounding context are recognized (cf. Halle et al. 1998).

A typological study of the posterior articulations

Chakir Zeroual
Centres des Etudes Universitaires de Taza and UMR 7018 CNRS / Sorbonne-Nouvelle

In this study we provide a phonetic typology of back articulations based on results of several phonetic experiences concerning the back consonants of Moroccan Arabic (MA) and particularly one recent and exhaustive fiberscopic study. We put this phonetic typology in relation with some phonological processes at work in MA, Classical Arabic (CA) as well as in several dialects of Modern Arabic. We try to establish if differences at phonetic level may account for some of the differences in the pattern of segments under consideration.

In a fiberscopic study, back and oral consonants of the MA were pronounced in the following contexts [-iCi-], [-aCa-] [-uCu-], [CCa-] and [CCi-] both in attested word and non-sense items. On the basis of those data it was possible to identify four types of articulations produced in the pharyngeal and laryngeal cavities.

1. “Pharyngeal” : produced with a backward movement of the root of the tongue towards the posterior pharyngeal wall ; observed for uvulars and emphatics. 2. “Epiglotto-pharyngeal” : produced between the tip of the epiglottis and the posterior pharyngeal wall ; observed for two consonants genrally transcribed with API symbols for pharyngeal consonants. These same consonants have a second articulation wich is 3. “aryepiglottal” produced between the base of the epiglottis and the arytenoids. 4. “Laryngeal” : produced at the level the glottis and observed for continuant laryngeal consonant of MA.

These fibroscopic data show that MA has epiglottal (consonants involving essentially the activity of epiglottis) and not pharyngeal consonants. The aerodynamic data suggest that epiglotto-pharyngeal articulation is simply a consequence of the aryepiglottal articulation.

Uvulars and emphatics are produced in MA with a retraction of the root of the tongue without aryepiglotic constriction, which proves the independence of these two articulations. In favor of this deduction we can mention another phonological process at work in MA according to which MA epiglottal consonants are more like laryngeal than uvular consonants. Our fiberscopic data are not in accord with Jakobson’s (1957) proposal that pharyngeal consonant of Arabic are in fact pharyngealized laryngeals.

However, we provide some phonetics and phonological arguments showing that the arabic consonants, which are epiglottals in MA, would be produced in other dialects of modern Arabic and probably in CA as well with a combination of “pharyngeal” and “aryepiglottal” articulations. This is probably why for Catford (1977) Arabic has pharyngeal and not epiglottal consonants, and for Jakobson (1957) the pharyngeal consonant are pharyngealized laryngeal ones.

We provide also some other phonetic and phonological cues showing that uvular consonants of MA are different of corresponding consonants in CA. Precisely, we put forward a phonogical process of nasalization at work in the rules of Coranic Recitation which seems to show that these consonants must be represented without features under the “Place” node (see also Halle, 1995) as it the case with the corresponding consonants in MA.

References

Halle, M. (1995). Feature geometry and feature spreading. *Linguistic Inquiry* 26 (1): 1-46.

Jakobson, R. (1957). Mufaxxama, the « emphatic » phonemes of Arabic. In E. Pulgram (ed.) *Studies presented to Jushua Whatmough on his 60th Birthday*. Mouton, the Hague: 105-115.