

MODELING THE EVOLUTION OF VOWEL SYSTEMS IN PHYLOGENY, ONTOGENY, AND LANGUAGE SPECIATION

Andrew R. Plummer Mary E. Beckman

The Ohio State University, Columbus, OH, USA

`{plummer.321,beckman.2}@osu.edu`

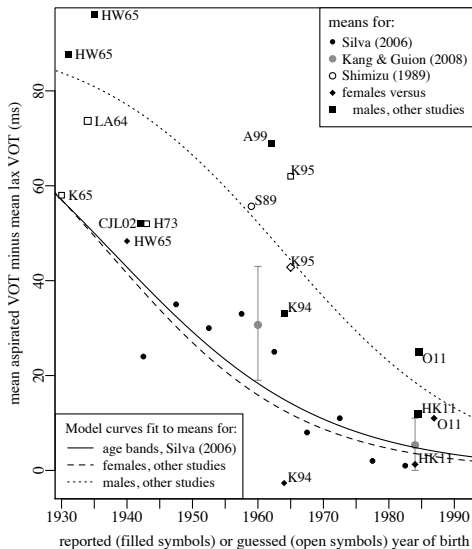
Jan. 16th, 2016

INCREMENTATION: CONCEPTS

LABOV (2007)

- ▶ Linguistic transmission is defined as the “unbroken sequence of native-language acquisition by children”
- ▶ “The continuity of dialects and languages across time is the result of the ability of children to replicate faithfully the form of the older generation’s language, in all of its structural detail,”
- ▶ “But linguistic descent can be preserved even when this replication is imperfect, that is, when language changes. This is the normal type of internal language change, termed CHANGE FROM BELOW...”
- ▶ “[I]nternal changes are generated by the process of INCREMENTATION, in which successive cohorts and generations of children advance the change beyond the level of their caretakers and role models, and in the same direction over many generations.”

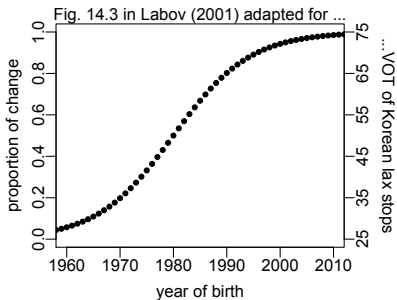
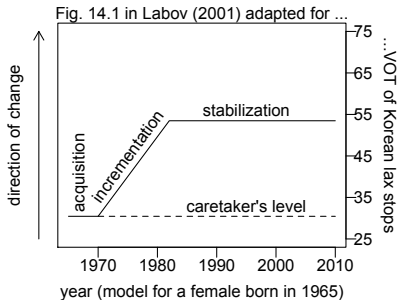
INCREMENTATION: CONCEPTS



INCREMENTATION: MODEL SKETCHES

LABOV (2007)

“Incrementation begins with the faithful transmission of the adult system, including variable elements with their linguistic and social constraints”

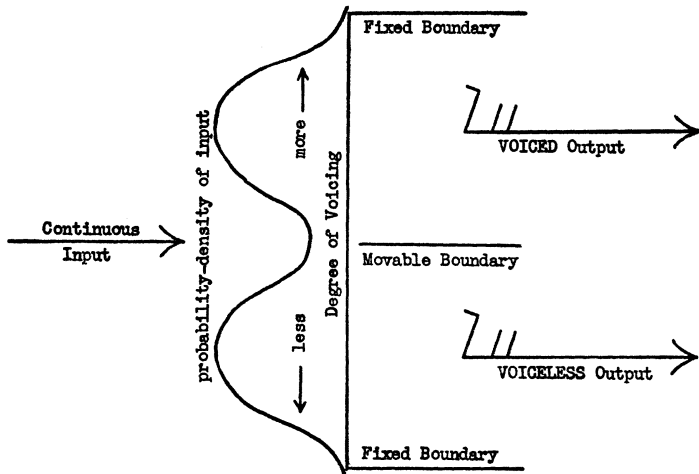


INCREMENTATION: MODEL SKETCHES

HOCKETT

- ▶ “The ‘space’ of all possible speech sound, either in articulation or acoustically, is a multidimensional continuum” (1965).
- ▶ “In any speech community, only certain DIFFERENCES of speech sound are functional. This breaks the continuous multidimensional space of all possible speech sound into a finite number of regions;”
- ▶ “Successive articulations aimed into the same region show a scatter, clustered around a LOCAL FREQUENCY MAXIMUM. The local frequency maxima are, or create, the ‘sort of conception of an average pronunciation’”
- ▶ “The frequency maxima for a particular language are then the ACOUSTIC ALLOPHONES of the language, and their projections on the axes are its DISTINCTIVE FEATURES.”

INCREMENTATION: MODEL SKETCHES



INCREMENTATION: MODEL SKETCHES

SOUND CHANGE

- ▶ “[A]s the density distribution varies for all the speakers of the community in pretty much the same way, the local maxima slowly wander about. This drifting of allophones, and hence of distinctive features, is SOUND CHANGE” (Hockett, 1965).
- ▶ This view has been adopted by sociolinguists since Weinreich et al. (1968) and applied extensively to vowel changes where the acoustic allophones are defined in terms of formant values.

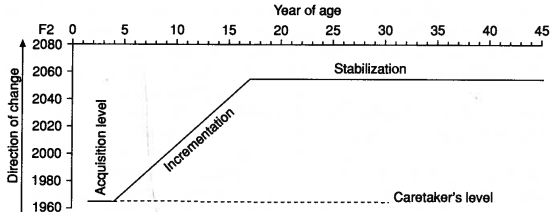


Figure 14.1 A linear model of incrementation for a single female speaker from 1 to 45 years of age

INCREMENTATION: MODEL SKETCHES

LABOV (1981), FIGURE 3.

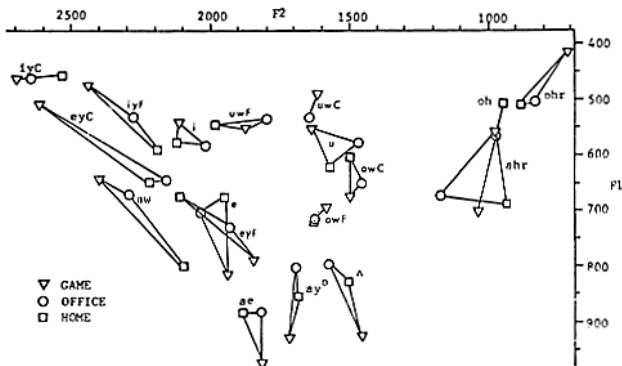


FIGURE 3.

Figure 3 shows the vowel system of Carol Meyers as analysed by Hindle. Each vowel shows not one but three means, for each of the three major social settings: the travel agency, dinner at home, and the bridge game. Free (uwF) is well front of center—not yet as far front as /i/, but well ahead of checked

INCREMENTATION: MODEL SKETCHES

LABOV (1990), FIGURE 6.

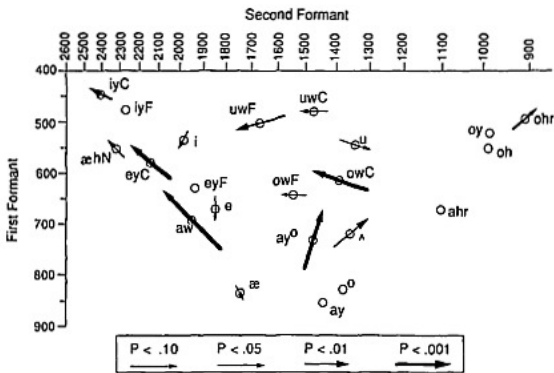
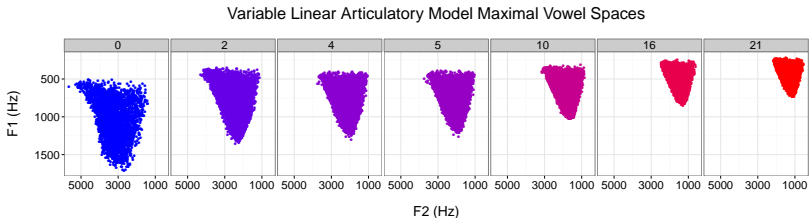


FIGURE 6. Movements of Philadelphia vowels in apparent time. Circles show mean values for 116 speakers in the Neighborhood Study. Vectors connect values for groups 25 years older & younger than the mean. F = free vowel; C = checked vowel; O = before voiceless finals.

INCREMENTATION: CONCEPTUAL DIFFICULTIES



INTER-SPEAKER VARIATION

- ▶ Inter-speaker differences dwarf even inter-phoneme differences, and they very much dwarf the kinds of subphonemic variation/drift that is being modeled.
- ▶ This means that vowel changes cannot be modeled in terms of acquisition of specific formant targets followed by incremental shifts of adolescents' formant values in the direction of change.

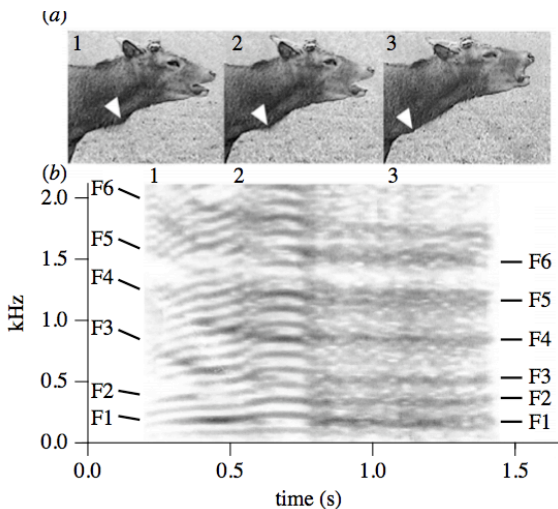
INCREMENTATION: CONCEPTUAL DIFFICULTIES

INTER-SPEAKER VARIATION

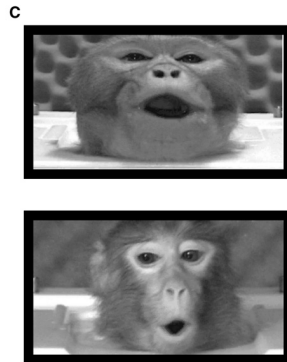
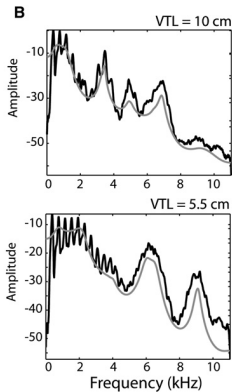
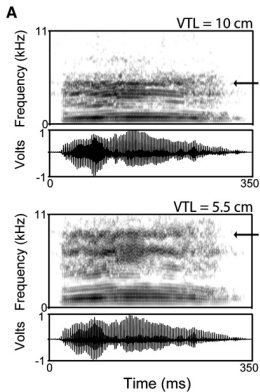
- ▶ Hindle (1978) distinguished between “technical solutions” to normalization and its “psychological” aspects.
- ▶ Previous work explicitly adopts some algorithm for technical normalization that either is or is like the algorithm that Don Hindle developed for Labov.
- ▶ “Supposing that gender differences are a mixture of the effects of vocal tract length differences and social factors, how could one separate the two types of influence, and arrive at a scaling factor that eliminated only the differences due to vocal tract length without removing the effect of social factors?” (Labov, 2006)

VOCAL TRACT LENGTH AS A SOCIAL FACTOR

But vocal tract length **is** a social factor, and not just in humans. Cf. Fitch & Reby (2001):

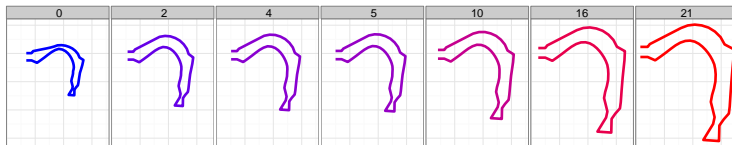


VOCAL TRACT LENGTH AS A SOCIAL FACTOR



MAXIMAL VOWEL SPACES AND STIMULI

Variable Linear Articulatory Model Midsagittal Representations



ARTICULATORY MODEL

- ▶ The VLAM (Boë & Maeda, 1998) is a computational model of the articulatory system and its speech production capacities.
- ▶ Midsagittal representations are wrought by configuring “articulatory blocks” (Lindblom & Sundberg, 1971; Maeda, 1990, 1991) corresponding to jaw height, tongue body position, tongue dorsum position, tongue apex position, lip protrusion, lip height, and larynx height.
- ▶ The VLAM is age-varying and capable of representing vocal tract lengths ranging from those of infants to young adults.

MAXIMAL VOWEL SPACES AND STIMULI

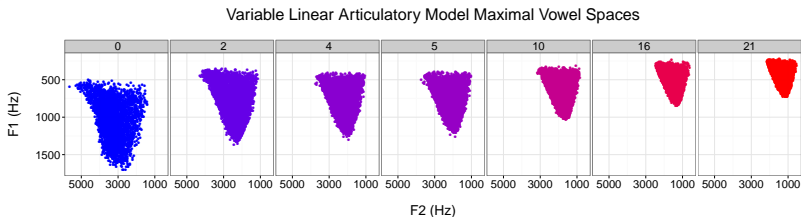


FIGURE : Maximal Formant Spaces (MFS) for each age.

MAXIMAL VOWEL SPACES

- ▶ Given an age in years, the set of all articulatory configurations of the VLAM at that age that do not result in occlusion of the oral cavity yield a corresponding **maximal vowel space** (MVS, Boë et al., 1989; Schwartz et al., 2007) for that age.
- ▶ We fix the following age index $AGES = \{0.5, 2, 4, 5, 10, 16, 21\}$ for indexing the MVSs discussed below. For each $a \in AGES$, let $MVS(a)$ be a dense sampling of the MVS for age a .

MAXIMAL VOWEL SPACES AND STIMULI

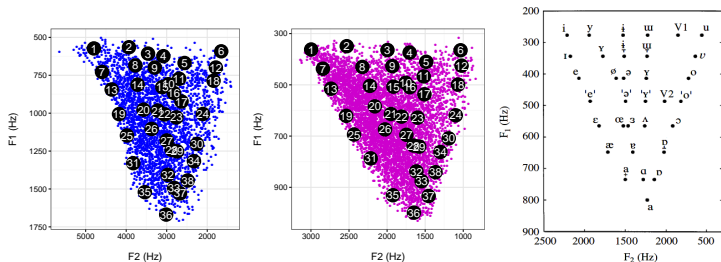
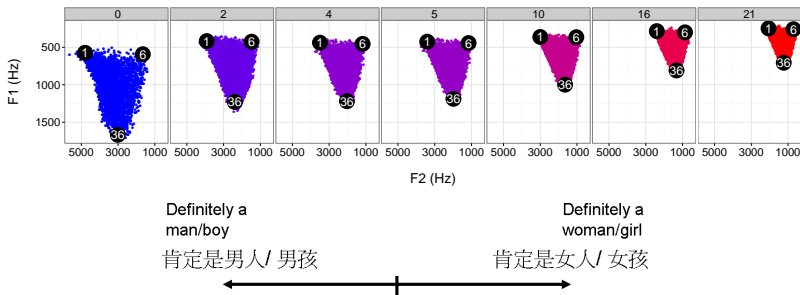


FIGURE : Prototype sets $P(0.5)$ (left), $P(10)$ (center), and those from Schwartz et al. (1997a).

あ い う
え お

RESPONSE DATA

Corner Vowel Stimuli within their Maximal Vowel Spaces

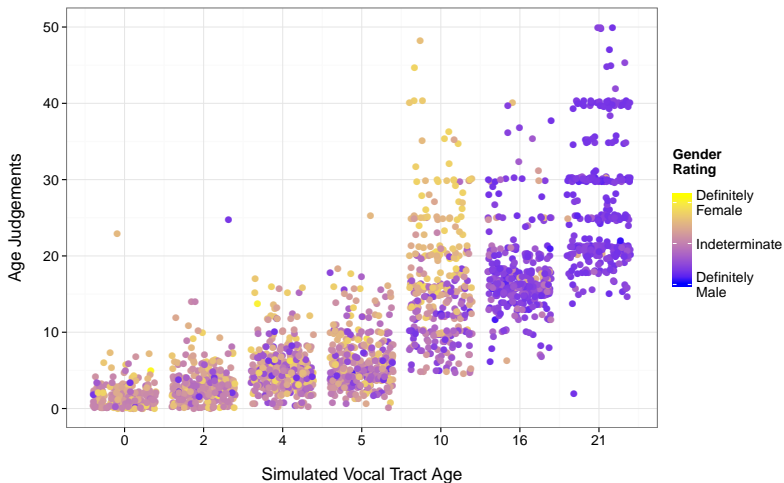


AGE AND GENDER RESPONSES

Cantonese-, American English- and Japanese-speaking listeners provided numeric age judgments and gender ratings along a visual analog scale ranging from "definitely male" to "definitely female" (or their Cantonese and Japanese equivalents) for subsets of the stimuli corresponding to the corner vowels /i/, /a/, and /u/.

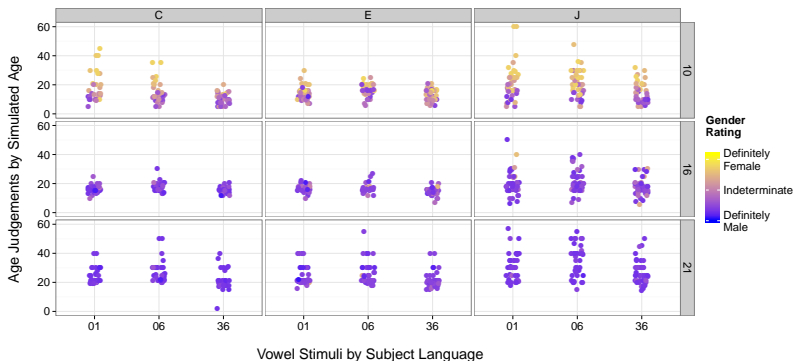
AGE AND GENDER DATA

Age Judgements vs. Simulated Vocal Tract Age (with Gender Ratings)



AGE AND GENDER DATA

Age Judgements vs. Simulated Vowel Stimuli (Across Languages)



This culture- and vowel- specific parsing of the age is in keeping with evidence that adult listeners' perception of talker information is more accurate in their own language than in an unfamiliar language (e.g., see Winters, Levi, and Pisoni, 2008, and references there)

FUNCTION AND IDENTITY INTERTWINED

Cheney and Seyfarth (1999): Playback experiment with combinations of dominant baboons' threat grunts and subordinate baboons' screams

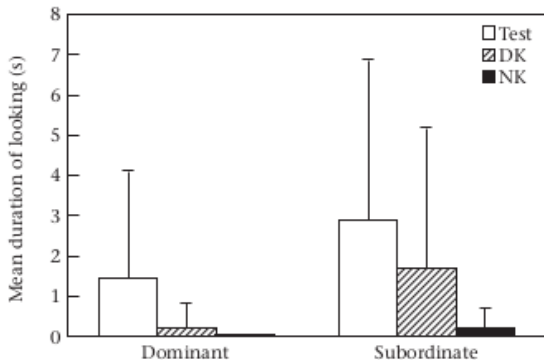


Figure 2. The mean duration that dominant and subordinate subjects looked towards each other following playback of sequences that included the calls of both of their close relatives (test condition), only the dominant subject's relative (DK condition), or neither of the subjects' relatives (NK condition).

FUNCTION AND IDENTITY INTERTWINED

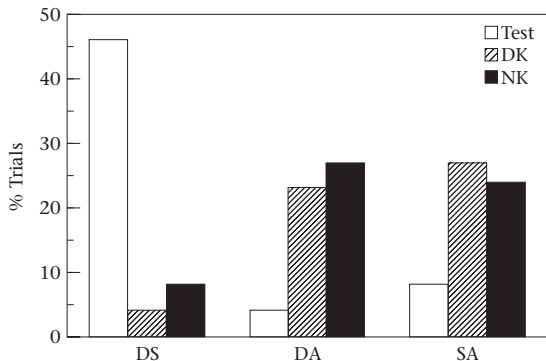


Figure 3. The proportion of subjects' first interactions with their partners that took various forms following each trial condition. Histograms show means for 26 dyads in the test, DK and NK conditions. DS: Dominant subject supplanted her subordinate partner; DA: dominant subject approached her subordinate partner without supplanting her or approached and interacted in a friendly manner with her subordinate partner; SA: subordinate subject approached her dominant partner or approached and interacted in a friendly manner with her dominant partner.

FUNCTION AND IDENTITY INTERTWINED



VOWEL CATEGORY ACQUISITION

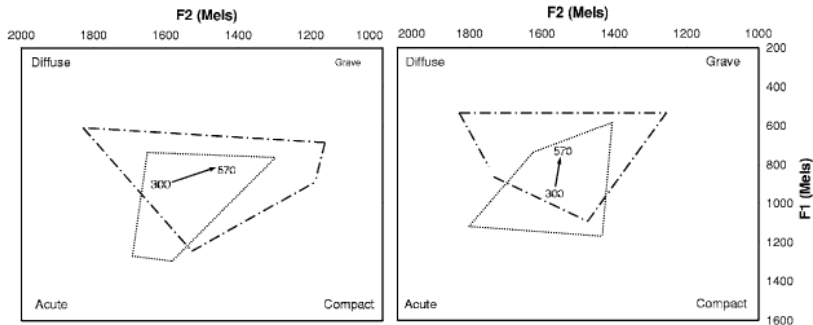


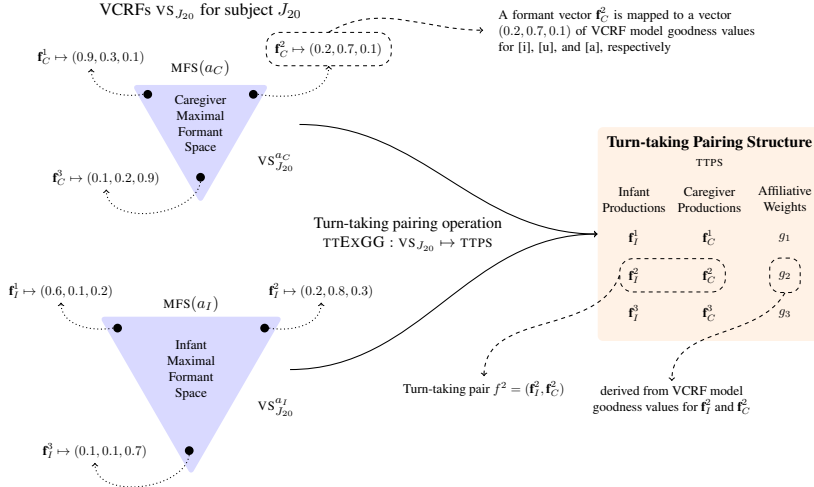
FIG. 4. Graphic summary of the findings for the English-learning infants (left) and French-learning infants (right). On both charts the arrow indicates the movement of the center of the vowel space as age increases from 300 to 570 days, the dotted-line quadrilaterals trace the periphery of the vowel space at 300 days, and the dashed-line quadrilaterals trace the periphery of the vowel space at 570 days.

These kinds of differences at 10 months are observable as early as 4 months (see Kuhl and Meltzoff (1996) for American English, and Ishizuka et al. (2007) for Japanese).

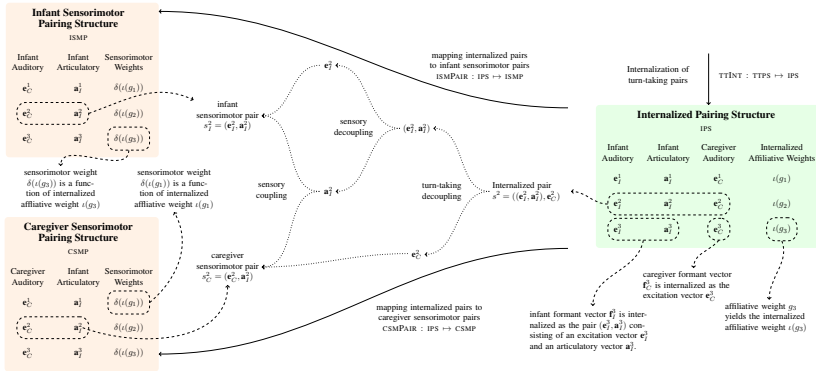
THE ACQUISITION OF VOWEL NORMALIZATION

Vowel Category Response Fields

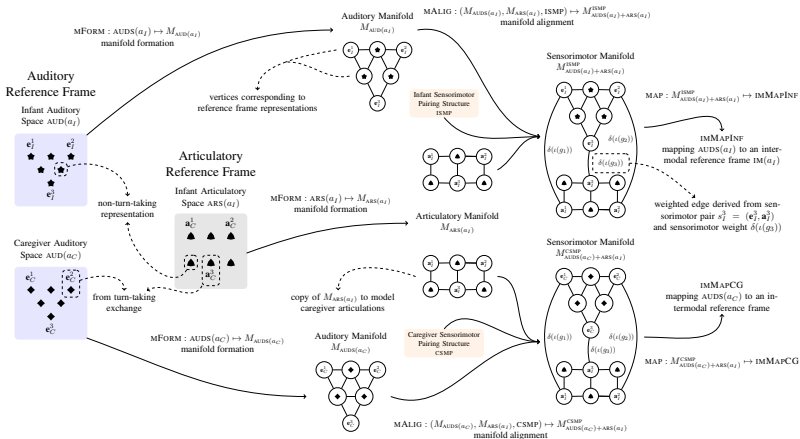
VCRFs vs J_{20} for subject J_{20}



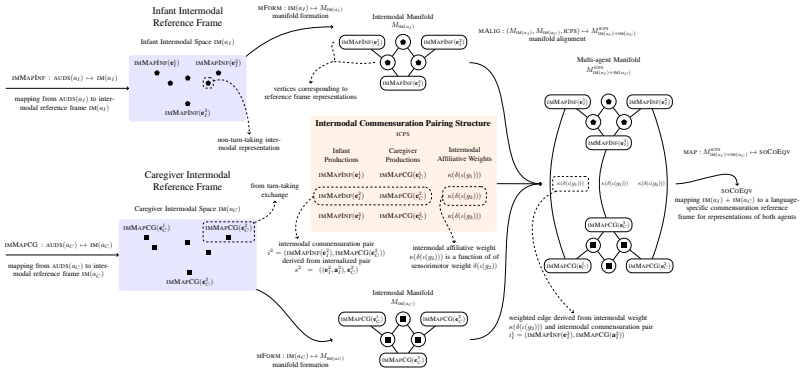
THE ACQUISITION OF VOWEL NORMALIZATION



THE ACQUISITION OF VOWEL NORMALIZATION



THE ACQUISITION OF VOWEL NORMALIZATION



ACKNOWLEDGEMENTS

Thanks to our collaborators Benjamin Munson and Lucie Ménard, to Catherine McBride for recruiting the Cantonese-speaking listeners, to Kiyoko Yoneyama for recruiting the Japanese-speaking listeners, and Chanelle Mays for recruiting the English-speaking listeners.

Work supported by NSF grants BCS 0729277 (to Munson) and BCS 0729306 (to Beckman).

- Boë, L., & Maeda, S. (1998). Modélisation de la croissance du conduit vocal. Espace vocalique des nouveaux-nés et des adultes. Conséquences pour l'ontogenèse et la phylogenèse. In *Journées d'Études Linguistiques: "La Voyelle dans Tous ces États"*, (pp. 98–105). Nantes, France.
- Boë, L., Perrier, P., Guérin, B., & Schwartz, J. (1989). Maximal vowel space. In *EUROSPEECH 09*, (pp. 281–284). Paris, France.
- Lindblom, J., & Sundberg, J. E. F. (1971). Acoustical consequences of lip, tongue, jaw, and larynx movement. *Journal of the Acoustical Society of America*, 50(4B), 1166–179.
- Maeda, S. (1990). Compensatory articulation during speech: Evidence from the analysis and synthesis of vocal-tract shapes using an articulatory model. In W. Hardcastle, & A. Marchal (Eds.) *Speech Production and Speech Modeling*, (pp. 131–149). The Netherlands: Kluwer Academic Publishers.
- Maeda, S. (1991). On articulatory and acoustic variabilities. *Journal of Phonetics*, 19, 321–331.
- Ménard, L., & Boë, L. (2000). Exploring vowel production strategies from infant to adult by means of articulatory inversion of formant data. In *Proceedings of the 6th International Conference on Spoken Language Processing (ICSLP 2000)*, (pp. 465–468). Beijing, China.

- Ménard, L., Schwartz, J., & Boë, L. (2002). Auditory normalization of french vowels synthesized by an articulatory model simulating growth from birth to adulthood. *Journal of the Acoustical Society of America*, 111(4), 1892–1905.
- Munson, B., Ménard, L., Beckman, M. E., Edwards, J., & Chung, H. (2010). Sensorimotor maps and vowel development in english, greek, and korean: A cross–linguistic perceptual categorization study (a). *Journal of the Acoustical Society of America*, 127, 2018.
- Okada, H. (1999). Japanese. In *Handbook of the International phonetic Association: a guide to the use of the International Phonetic Alphabet*, (pp. 117–119). Cambridge, U.K.; New York, New York: Cambridge University Press.
- Schwartz, J., Boë, L., & Abry, C. (2007). Linking dispersion–focalization theory and the maximum utilization of the available distinctive features principle in a perception–for–action–control theory. In M. . J. Sole, P. S. Beddor, & M. Ohala (Eds.) *Experimental Approaches to Phonology*, (pp. 104–124). Oxford University Press.
- Schwartz, J., Boë, L., Vallée, N., & Abry, C. (1997a). The dispersion–focalization theory of vowel systems. *Journal of Phonetics*, 25, 255–286.

- Schwartz, J., Boë, L., Vallée, N., & Abry, C. (1997b). Major trends in vowel system inventories. *Journal of Phonetics*, 25, 233–253.
- Vallée, N., Boë, L., & Payan, Y. (1995). Vowel prototypes for upsids's 33 phonemes. In *Proceedings of ICPhS 2*, (pp. 424–427). Stockholm.