

WW29. Labial position and acoustics of Korean and English high vowels. Sun-Ah Jun (Dept. of Linguistics, Ohio State Univ., 204 Cunz Hall, Columbus, OH 43210)

This paper examines and contrasts the labial configuration and formant frequencies of Korean and English high vowels. Korean has three high vowels, /i, i, u/, and English has four, /i, I, U, u/. The lip gestures and formant frequencies were compared within each language and across language to determine whether the idea of maximal dispersion [Liljencrants and Lindblom, 1972] can account for the formant frequencies and also be extended to account for the labial configurations. Three male speakers of each language produced each vowel in four different contexts with five repetitions. The production of each word was videotaped and the sound wave simultaneously recorded. The lip configurations were assessed using measurements similar to those in Linker [UCLA Working Papers in Phonetics 56 (1982)]. Formant frequencies were measured using an LPC formant tracker. Preliminary results suggest that Korean high vowels can be grouped into two categories, /i, i/ vs /u/, on the basis of measurements of lip position. The English vowels similarly can be grouped into three categories, /i, I/ vs /U/ vs /u/. For the vowel /u/, Korean and English used different degrees of lip protrusion and, for the vowel /i/, they differed in the relationship between the horizontal width and the vertical opening. The acoustic spaces of Korean and English high vowels both support the dispersion theory in covering the same amount of the vowel space, although Korean high vowels all had lower *F*₂ values.

WW30. Variability of motor strategies. Michèle Gentil (CHU Pitié Salpêtrière, URA CNRS 385, Lab. Physiologie de la Motricité, 91 bd de l'hôpital, 75634 Paris cédex 13, France)

The activity patterns of the mandibular and labial muscle systems, respectively, were investigated for the purpose of determining whether the functions of the jaw and lips are based on a set of universal neuromotor rules. In the first experiment, intramuscular wire electrodes were placed into the eight muscles of the jaw. The activity of these muscles was recorded along with the displacement of the jaw in two-dimensional space, from three American subjects, during the production of speech at different rates. In the second experiment, intramuscular wire electrodes were placed into seven perioral muscles. EMG activity was recorded along with two-dimensional upper and lower lip movements from three French subjects during labial closure associated with speech production. Results indicated that different individuals use individual motor strategies for the production of the jaw and lip movements, and that the same motor objective can be accomplished by the nervous system in many different ways.

WW31. Lip rounding in Amoy and Mandarin high vowels: Maximum dispersion, or adequate separation. Ho-Hsien Pan (Dept. of Linguistics, Ohio State Univ., Columbus, OH 43210)

There are two hypotheses about the relationship between phonological contrasts and phonetic feature scales. Some phoneticians propose that values are chosen so that contrasting phonemes are maximally separated [e.g., Liljencrants and Lindblom, *Language* 48, 840-862 (1972)], while others claim that they need only to be adequately separated [e.g., Maddieson, *UCLA Working Papers in Phonetics* 36, 49-69 (1977)]. This paper tests the competing hypotheses by comparing lip position in Mandarin [i], [y], [u] and Amoy [i] and [u]. According to the maximum dispersion hypothesis, the degree of roundness of [i] and [u] in both languages should be the same. According to adequate separation, on the other hand, the lip position of Mandarin [i] will be more spread than [i] in Amoy, and Mandarin [u] will be more rounded, since there are three high vowels in Mandarin but only two in Amoy. Amoy and Mandarin data were collected from three bilingual speakers to ensure that differences in roundness were due to language differences, not personal characteristics. Preliminary results show that the lip position of [i] in both languages is the same, while the [u] in Mandarin is more rounded than Amoy [u]. Both hypotheses seem to be true here.

WW32. Comparison of lip rounding in German and English vowels. Monica Crabtree and Claudia Kurz (Dept. of Linguistics, Ohio State Univ., Columbus, OH 43210-1229)

Lip displacements in German and English vowels were compared in order to test two claims; first, that German rounded vowels are more rounded than their English counterparts [Disner, *UCLA Working Papers in Phonetics* No. 58 (1983)] and, second, that [y] is less rounded than [u] in German [S. Wood, *J. Acoust. Soc. Am.* 80, 391-401 (1986)]. Disner's claim is related to the idea that the number of contrasting vowels in a language affects the dispersion within the vowel space [Keating and Huffman, *Phonetica* 41, 191-207 (1984)]. The number of vowels may also constrain contextual variation due to consonantal environment in different ways in the two languages. This was also tested. Three native speakers of each language were videotaped saying four tokens of each vowel in their respective languages, in bilabial-alveolar context and in velar-alveolar context. Various measurements of upper and lower lip position were analyzed using ANOVA. Preliminary results indicate that German [u, u, o, o] are indeed more rounded than English [u, u, o, o]. However, the results with respect to [y] do not support Wood's claim.

THURSDAY AFTERNOON, 24 MAY 1990

104 KELLER, 1:00 TO 4:00 P.M.

Session XX. Speech Communication X: Cross Modal and Auditory Speech Perception

Philip F. Seitz, Chairman

Center for Auditory and Speech Sciences, Gallaudet University, Washington, DC 20002

Contributed Papers

1:00

XX1. Lipreading with vibrotactile vocoders. Lynne E. Bernstein (Ctr. for Auditory and Speech Sci., Gallaudet Univ., 800 Florida Ave., N.E., Washington, DC 20002), Marilyn E. Demorest (Univ. of Maryland, Baltimore County, Catonsville, MD 21228), Michael P. O'Connell

(Central Inst. for the Deaf, St. Louis, MO 63110), and David C. Coulter (Ctr. for Auditory and Speech Sci., Gallaudet Univ., Washington, DC 20002)

A training study was conducted to compare aided lipreading of normal-hearing and deaf adults, each assigned to one of three vibrotactile

vocoders. Vocoders were (1) the Queen's University/Central Institute for the Deaf (QU/CID) vocoder, with one-third octave filter spacing and logarithmic output scaling; (2) the QU/CID vocoder with linear output scaling; and (3) the GU (Gallaudet University) vocoder designed for greater resolution than the others in the *F2* region and linear output scaling. Subjects received stimuli in baseline (no vocoder) and treatment (vocoder) conditions. In addition, two subjects served as visual-only controls. Stimuli were provided by a live talker and two talkers prerecorded on laser videodisc (Bernstein and Eberhardt, 1986). Preliminary analysis of the results suggests that (1) the QU/CID linear vocoder was most effective, followed by the GU vocoder with linear output; and (2) regardless of experimental condition, normal-hearing subjects' lipreading improved over the approximately 65-h experiment. Results with deaf adults, along with results of the visual control subjects, suggest that careful control of visual learning is needed in experiments involving aided lipreading. Results will be compared with a previous study that involved similar procedures and several different transformations of fundamental frequency for a single vibrotactile channel. [Research supported by NIH.]

1:15

XX2. The influence of orthographic information on the identification of an auditory speech event. Jody K. Layer, Richard E. Pastore, and Ellen Rettberg (Dept. of Psychol., SUNY, Binghamton, NY 13901)

It has been shown that the identification of an auditory or visual speech event can be influenced by information from the other modality when the information is perceived to arise from the same event. Employing a selective attention task, Logan *et al.* (1990) demonstrated an influence of an auditory speech event on the identification of orthographic characters over a range of stimulus onset differences. The current research investigates the influence of orthographic information on the identification of the initial phoneme of an auditory speech event. The stimuli were nonword CVC syllables. The stimuli were edited natural speech and orthographic representations of these stimuli plus a set that was neutral with respect to the auditory set. A range of stimulus onset differences was employed. The results show that when the auditory and visual information agree, there is faster responding for the identification of the initial auditory phoneme. When the information is discrepant, responding is slowed. These results do not support a qualitative change in perception with differing information in the two modalities. Implications for the nature of cross-modal integration and speech event processing will be discussed.

1:30

XX3. Exploring the basis of the "McGurk effect": Can perceivers combine information from a female face and a male voice? Kerry P. Green (Dept. of Psychol., Univ. of Arizona, Tucson, AZ 85721), Erica B. Stevens, Patricia K. Kuhl, and Andrew M. Meltzoff (Univ. of Washington, Seattle, WA 98185)

In the "McGurk" effect, observers typically report the illusory syllable /da/ when they hear the auditory syllable /ba/ presented in synchrony with a video display of a talker saying /ga/. In such experiments, there is usually congruence between the two modalities in that the same talker produces both the auditory and the visual signals. In the experiments reported here, the effect of reducing the congruence between the two modalities on the magnitude of the McGurk effect was examined. This was accomplished by dubbing a male talker's voice onto a video tape containing a female talker's face, and a female talker's voice onto a video tape containing a male talker's face. These "cross-dubbed" video tapes were compared to normal video tapes in which the male talker's voice was dubbed onto a male talker's face, and the female talker's voice was dubbed onto a female talker's face. The results show that even though there was clear incompatibility in the talker characteristics between the auditory and visual signals for the cross-dubbed stimuli, there was little difference in the magnitude of the effect compared to the normal stimuli. These

results indicate that the mechanism for integrating speech information from the two modalities is not sensitive to certain incompatibilities, even when they are perceptually apparent. [Work supported by NIH.]

1:45

XX4. Cross-modal semantic priming of neighbors of multisyllabic words. Paul A. Luce and Michael S. Cluff (Language Perception Lab., Dept. of Psychol., SUNY, Park Hall, Buffalo, NY 14260)

The present experiment was designed to test the principle of delayed commitment in spoken word recognition by using a cross-modal priming paradigm. Subjects were asked to make a lexical decision on visually presented targets that were preceded by auditory primes. These auditory primes consisted of spondee (words containing two individual lexical items) that had second syllables with at least two meanings. For example, the second syllable of "baseball" could refer to either a round object for throwing or to a formal dance. Targets were words that were either related to the alternate meaning of the prime's second syllable or were unrelated. For example, after hearing "logjam," subjects were presented with either "JELLY," which is related to an alternate meaning of "jam," or "BOMB," which is unrelated to either meaning of "jam." Evidence of priming in the related condition suggests that multiple candidates for recognition remain activated until well after the word's isolation point, contrary to the predictions of cohort theory. However, consistent with the neighborhood activation model, these results demonstrate that word recognition operates by a principle of delayed commitment.

2:00

XX5. Effects of visual word stimuli on speech perception. Tadahisa Kondo and Kazuhiko Takehi (NTT Basic Res. Lab., 3-9-11, Midori-cho, Musashino-shi, Tokyo, 180 Japan)

The effects of visual word stimuli on speech perception in Japanese are investigated. An experiment is constructed to discriminate effects mediated by phonological and semantical codes of written words. Subjects are simultaneously presented with a visual word stimulus and an auditory word stimulus embedded in a short sentence. The recognition rates for a syllable in a spoken word, where the syllable is either replaced with white noise or has white noise added to it, are measured in various conditions in which various types of visual word stimuli are presented. Four types of visual word stimuli are used: (1) matching words, (2) nonword strings with the same pronunciation, (3) associative meaning words, and (4) nonrelative words in regard to auditory word stimuli. Additionally, non-character stimuli are used as a control condition. The mean syllable recognition rates for five conditions are in order of (1), (2), (3), control, and (4), from higher ones. The rates for (1) and (2) are significantly different from the rate for the control condition. Consequently, both phonological and semantical codes of visual words are important factors affecting speech perception.

2:15

XX6. Similarity neighborhoods of spoken two syllable words: Retroactive effects on multiple activation. Michael S. Cluff and Paul A. Luce (Language Perception Lab., Dept. of Psychol., SUNY, Park Hall, Buffalo, NY 14260)

This research examined the recognition of two-syllable spoken words and the means by which the auditory word recognition system deals with ambiguous stimulus information. The perceptual identification of two-syllable words comprised of two monosyllabic words (spondee) was examined. Individual syllables within a spondee were characterized as either "easy" or "hard" depending on the neighborhood characteristics of the