Contributed Papers


Studies of the effect of speaking rate on speech segment durations reveal that the VOTs of long lag and prevoiced consonants vary directly with speaking rate, whereas those of short lag consonants remain relatively constant across speaking rates (i.e., Kessinger and Blumstein, 1998; Miller and Volaitis, 1989). Koreans have three voiceless stop consonants at each place of articulation. These stops vary in VOT with values in the long lag and short lag ranges. These phonemes are also marked by differing fundamental frequency patterns. The purpose of this study was to investigate the effects of different speaking rates on the acoustic parameters of the Korean voiceless stop consonants. Sixteen Korean adults (8 male and 8 female) repeated productions of the lax, reinforced, and aspirated forms of /p/, /t/, and /k/ with the vowel /i/. These repetitions were made of the isolated CV syllables and of the syllables in a Korean sentence context. The syllables were analyzed for segment durations of hold, transition and vocalic portions and fundamental frequency contours. The speech was recorded using a DAT recorder and analyzed using a CSL 4300B system. The effect of the speaking rate on the different phonemes will be discussed [R. Kessinger and S. Blumstein, “Effects of speaking rate on voice-onset time and vowel production: Some implications for perception studies,” J. Phonetics 26, 117–128 (1998); J. Miller and L. Volaitis, “Effect of speaking rate on the perceptual structure of a phonetic category,” Percept. Psychophys. 46, 505–512 (1989)].

4aSC2. Speaking rates of American and New Zealand varieties of English. Michael Robb (Univ. of Connecticut, Dept. of Commun. Sci., 850 Bolton Rd., Storrs, CT 06269), Yang Chen (Univ. of Wyoming, Dept. of Commun. Disord., Laramie, WY), and Margaret Macalagan (Univ. of Canterbury, Dept. of Speech–Lang. Therapy, Christchurch, New Zealand, m.macalagan@sph.canterbury.ac.nz)

Measures of speaking rate provide an index of speech timing control. The present study sought to evaluate speech timing control in two varieties of English, American English (AE), and New Zealand English (NZE). Measures of speaking rate were calculated for 40 adult AE speakers which were then compared to a group of 40 adult NZE speakers. Results of the analysis identified significantly faster speaking rate and articulation rate among the NZE speakers compared to AE speakers. Vowel changes occurring in both varieties of English over the past decades, as well as contrasting rhythmic properties, are proposed to account for the rate differences. Because temporal differences can be a component of disordered speech and are a major component in certain motor speech disorders and disorders of fluency, normative speaking rate data are often utilized by speech-language pathologists. The present results, viewed from a clinical standpoint, would suggest that the particular variety of English spoken by a client is an important variable when addressing disorders of speech timing.

4aSC3. Effects of speaking rate on temporal patterns of English. Bruce L. Smith (Dept. of Commun. Sci. and Disord., Northwestern Univ., 2299 N. Campus Dr., Evanston, IL 60208 b-smith2@northwestern.edu)

Previous research [B. L. Smith, J. Acoust. Soc. Am. 108, 2438–2442 (2000)] has indicated that individual subjects show considerable variation in the extent to which they manifest a number of temporal patterns in their speech (e.g., vowel lengthening before voiced obstruents, final-syllable vowel lengthening, etc.). At least in part, some of the variation across speakers was found to be related to the fact that some subjects naturally tend to talk somewhat faster/slower than others. The present study extended these previous findings concerning variations in temporal performance related to rate of speech by having 15 subjects produce several different CVC target stimuli in sentences spoken at both normal and fast rates. The temporal patterns observed as a function of these experimental manipulations of speaking rate generally supported the previous findings related to natural variations in rate of speech across subjects. It was found, for example, that vowel lengthening preceding voiced obstruents tended to be less when subjects spoke at a fast versus a normal rate. In contrast, phrase-final vowel lengthening was typically greater when subjects spoke at a faster rate. Possible reasons for these different vowel-lengthening effects may relate to the temporal domain that influences them.

4aSC4. Exploring underlying pitch targets in English statements. Yi Xu and Chun X. Xu (Dept. of Commun. Sci. and Disord., Northwestern Univ., Evanston, IL 60208)

This study investigates F0 contours and their potential association with underlying pitch targets in English statements with and without narrow focus. Eight native speakers of American English read aloud 21 sentences at 2 speaking rates with 7 repetitions. The initial, middle, and final words in these sentences vary in focus status (focus or no focus) and word length (monosyllabic, disyllabic, or trisyllabic). F0 curves are extracted using a method that combines automatic vocal cycle detection and manual rectification. Preliminary results show that F0 peaks occur earlier when the vowel of the accented syllable is phonologically long than when it is phonologically short. This seems to suggest that a nonfinal pitch accent in an English statement is probably associated with a high static target. Preliminary analyses also show that F0 contours of mon accented syllables are influenced more by the preceding accent than by the following accent. Furthermore, the influence of the accented syllable on the F0 of the fol-
loving (but not the preceding) nonaccented syllable is found to reduce over time. These patterns seem to suggest that inter-accent FO contours in English probably also result from certain underlying pitch targets rather than from linear or nonlinear interpolation between adjacent pitch accents.

4aSC5. Contrastive focus in Beijing Mandarin. Yiya Chen (Linguist. Dept., SUNY-Stony Brook, Stony Brook, NY 11794-4367)

Contrastive focus in Beijing Mandarin. This study investigated the acoustic realization of contrastive focus in Beijing Mandarin. Here a contrastive focus is defined as a mechanism to signal contrast or make corrections in discourse. Two questions were asked. First, what are the acoustic cues employed to convey contrastive focus? Second, what are the linguistic constraints on the manifestation of prosodic cues? Four different renditions of each test sentence were elicited from the subjects with contrastive focus on different words to convey different pragmatic meanings. Preliminary results suggest: (1) All three acoustic cues under investigation FO, overall intensity, and duration are employed, with different degrees of consistency, to convey contrastive focus; (2) the manifestation of these cues also appear to be context sensitive. The implications of these phonetic facts for possible phonological organization of contrastive focus in Beijing Mandarin will be discussed. [Work supported by NSF Grant No. SBR 9600930 to Marie Huffman.]

4aSC6. Toward a perception based model of the production of prosody. Gazeoog Dogil and Bernd Moebius (Inst. of Natural Lang. Processing, Univ. of Stuttgart, Azenbergstr. 12, D-70174 Stuttgart, Germany)

The speech production model proposed by Guenther, Perkell, and co-worker’s (Guenther et al., Psych. Rev. 15 (1998); Perkell et al., J. Phon. 28 (2000)] is generalized from the predominantly segmental perspective to a new theory of the production of prosody. It is posited that the only invariant targets of the production of prosodic events are perceptual targets, characterized as multidimensional regions in the perceptual space. Prosodic gestures are planned as trajectories that reach and traverse the perceptual target regions. Successfully executed gestures produce acoustic realizations of perceptually relevant prosodic events. The prosodic interpretation of the speech production model is structured around a hierarchy of prosodic domains: discourse structure, information structure, and accentual structure. It is assumed that the linguistically relevant and phonologically distinctive functions of prosodic features are represented by internal models. Once acquired, the phonemic settings pertinent to these models are stable and resistant to change, and they do not rely on continuous auditory feedback. We further suggest that the mechanisms involved in prosody control, may not differ categorically from those that control segmental speech production.

4aSC7. Preliminary studies on respiratory activity in speech. Peter Ladefoged (Dept. of Linguist., UCLA, Los Angeles, CA 90095-1543, oldfogey@ucla.edu) and Gerald Loeb (Biomed. Eng., USC, Los Angeles, CA 90089-1112)

Work at Edinburgh University in the 1950s, summarized by Ladefoged [Three Area of Experimental Phonetics (Oxford University Press, Oxford, 1967)] claimed that (1) the mean power for speech was provided by actions of the respiratory muscles. The external intercostals check the outgoing airflow when speaking with high lung volumes, and the internal intercostals, rectus abdominis, and other muscles become more active as lung volume decreases. (2) Stressed syllables are produced by increases of respiratory power above the mean, usually by further activity of the internal intercostals. Hixon and Weismer (1995) pointed out technical limitations of these early studies, but wrongly indicated that they were of little value. These studies have been replicated with different methods to measure intramuscular EMG and airflow and pressure. The two subjects studied to date follow the pattern previously described. In addition the improved techniques indicate that speakers exercise further control by bringing in stabilizing antagonistic muscles. These results are consistent with the general hypothesis that the central nervous system is able to control subglottal pressure by using various combinations of passive elastic and active muscle forces, thereby achieving motor equivalence and acoustic invariance over a wide range of lung volumes.


In Spanish /b, d, g/ are usually realized as voiced approximants in all syllabic contexts unless following a stop—a phenomenon called spirantization. However, in North-Central Peninsular Spanish (NCS) voiceless fricatives rather than voiced approximants are produced. In the majority of cases this occurs in stressed syllables. This study examines whether stress is a factor in the likelihood of frication and devoicing of coda /b, d, g/ in this dialect. Nine native speakers of NCS were recorded producing nonce words with /b, d, g/ in coda position in both stressed and unstressed syllables. Measurements were made of vowel and consonant duration, presence and absence of frication and voicing, and voicing duration. The results indicate that frication is more likely in stressed syllables than in unstressed syllables. This suggests that in stressed syllables a higher subglottal pressure produces higher airflow across the glottis, favoring frication. In turn, frication inhibits voicing due to conflicting aerodynamic requirements between the two. It is concluded that stress is a factor in spirantization and that it may indirectly affect the voicing properties of /b, d, g/. [Work supported by Del Amo and the Basque Government.]


Native-language phonemes combined in a non-native way can be misperceived so as to conform to native phonotactics; e.g., English listeners are biased to hear syllable-initial [tr] rather than the illegal [tl] (Massaro and Cohen, 1983; Pitt, 1998). What sort of linguistic knowledge causes phonotactic perceptual bias? Two classes of models were compared: unit models, which attribute bias to the listeners differing experience of each cluster (such as their different frequencies); and structure models, which use abstract phonological generalizations (such as a ban on [coronal] sequences). Listeners (N = 16 in each experiment) judged synthetic 6 × 6 arrays of stop-sonorant clusters in which both consonants were ambiguous. The effect of the stop judgment on the log odds ratio of the sonorant judgment was assessed separately for each stimulus token to provide a stimulus-independent measure of bias. Experiment 1 compared perceptual bias against the onsets [bw] and [dl], which violate different structural constraints but are both of zero frequency. Experiment 2 compared bias against [dl] in CCV and VCCV contexts, to investigate the interaction of syllabification with segmentism and to rule out a compensation-for-coarticulation account of Experiment 1. Results of both experiments favor the structure models. [Work supported by NSF IGERT.]

4aSC10. The influences of phonotactic constraints of words on speech production. Patricia Ward and Jan Charles-Luce (Univ. at Buffalo, 122 Cary Hall, 3435 Main St., Buffalo, NY 14214)

It is known that words are stored in the mental lexicon with certain lexical and phonological properties and that these properties have consequences for the perception and production of words. Phonotactic probability is a phonological property of words. It is defined by segment position frequency and bi-phone co-occurrence frequency. A word can have a high probability pattern or a low probability pattern. In this study, the effects of the phonotactic probabilities of words on speech production have been examined in children with typical speech and language skills and children with phonological disorders. Typical children between the ages of 4–6,