Stimulus range as a determinant of phoneme boundaries along synthetic consonant continua

Michael Studdert-Kennedy

Citation: The Journal of the Acoustical Society of America **60**, S92 (1976); doi: 10.1121/1.2003610
View online: https://doi.org/10.1121/1.2003610
View Table of Contents: https://asa.scitation.org/toc/jas/60/S1
Published by the Acoustical Society of America

study was designed to examine vowel duration as a consonant voicing cue in naturally produced speech. Several CVC's were recorded by both authors as stressed citation forms. Vowel durations were measured and parts were removed from: the endmost section of the vowel, the middle of the vowel, or the end *and* the middle of the vowel. At least for these stimuli produced by these speakers, no clear relation is seen between the length of the preceding vowel and the perception of the voicing characteristic of the final consonant.

3:50

LL12. Perception of vowel duration. William S-Y. Wang (Project on Linguistic Analysis, 2222 Piedmont Ave., University of California, Berkeley, CA 94720), Ilse Lehiste (Department of Linguistics, Ohio State University, Columbus, OH 43210), Chin-Kuang Chuang, and Nancy Darnovsky (Project on Linguistic Analysis, 2222 Piedmont Ave., University of California, Berkeley, CA 94720)

Lehiste and Pisoni have independently shown that the perception of vowel duration is influenced by the F_0 pattern that accompanies the vowel. Here we report some related results. Sixteen subjects are asked to judge the relative duration of pairs of synthesized stimuli. The stimulus pairs are made up of the vowels /i/, /a/, /a^i/, and nonspeech, where nonspeech is a single formant at 1500 Hz. The F_0 patterns are: level = 120 Hz, rising = 105—135 Hz, and falling = 135 to 105 Hz. The durations are 220, 240, 260, and 280 msec. The orthogonal combination of these factors makes up a total of 512 stimulus pairs. In comparing duration judgements of stimuli of rising (R), falling (F), and level (L) F_0's, we found that the following relation obtains: R > F > L, where " > " denotes "perceived as having greater duration." This relation holds to the same extent for both speech and nonspeech stimuli. There is no signficant difference in the perceived duration between /i/ and /a^i/. The most consistent effect, found also earlier in a pilot experiment we performed with 58 subjects, is that /i/ is heard as longer than /a/, in a ratio of 60% : 40%. Since /i/ is normally shorter than /a/ in natural speech, we posit a compensation mechanism mediating between perception and production that is responsible for this negative correlation between vowel height and duration. It is interesting to note that although the processes of duration judgment and pitch judgement (cf. the abstract by Chuang and Wang) are different, they seem to share a common compensation mechanism between production and perception. [Work supported by NSF Grant No. BNS 76-00017-Wang-2/78J34.2.]

4:00

LL13. Influence of vowel height, intensity, and temporal order on pitch perception. Chiu-Kuang Chuang and William S-Y. Wang (Project on Linguistic Analysis, 2222 Piedmont Ave., University of California, Berkeley, CA 94720)

Recently, it has been observed that high vowels are produced with a higher F_0, even in tone languages. Earlier, H. Fletcher reported that louder pure tones are perceived as having a lower pitch, but only in the range 50 to 500 Hz, which is the range of F_0 in speech. Here we report some experiments which investigate these relations. Fifteen subjects are asked to judge the relative pitch of pairs of synthesized vowels. The vowels are /i, e, u, a/ with intensity differences (ΔI) of 0, 10, 20, and 30 dB SPL, and at frequency differences (ΔF_0) of 0, 1, 2, and 3 Hz, with a base frequency of 100 Hz. Each vowel is 300 msec long and the inter stimulus interval is 400 msec. The orthogonal combination of these factors, i.e., four vowel types, four ΔI's, four ΔF_0's, and the two positions within each pair, makes up a total of 512 stimulus pairs. The interaction between pitch perception and vowel height is highly significant statistically. Pooled across all subjects, /a/ is heard as 2.2 Hz higher than /u/, 0.8 Hz higher than /i/, and 0.2 Hz higher than /e/. The intensity factor by itself appears to have no effect on pitch perception, but interacts with the postion factor in influencing pitch perception. The negative correlation between production and perception in the relation between vowel height and pitch is interpreted as the result of a compensation mechanism mediating between production and perception. [Work supported by NSF Grant No. BNS 76-00017-Wang-2/78J34.2.]

4:10

LL14. An acoustic correlate of syllabicity in english. H. Semiloff (Speech Communications Research Laboratory, Santa Barbara, CA 93109)

The purpose of this research was to test three hypotheses (1) Increasing the duration of a nonsyllabic segment will cause the segment to be perceived as syllabic. (2) The perception of syllabicity is categorical. (3) Listeners will differ in their labeling responses to stimuli when told to judge them in different speech styles. Seven words were synthesized: "blow," "plight," "dress," "crest," "prayed," "broke," and "sport." The steady-state portion of the /l/ or the /r/ (for "sport," aspiration of /p/) was lengthened in 10-msec increments. Subjects were asked to decide whether the word was monosyllabic or disyllabic: "blow-below," "plight-polite," "dress-duress," "crest-caressed," "prayed-parade," "broke-baroque," and "sport-support." Five groups of 15 listeners each participated in the labeling task. Four groups heard the stimuli with a precursor frame, "The word you will hear next is ___" in one of four speech styles: formal-slow, formal-fast, casual-slow, casual-fast. Each group was told to judge the words according to the criterion of being spoken in one of the designated styles. The fifth group heard the stimuli with no precursor and received no style instructions. Results showed that (1) increased durations of /r/ and /l/ did result in perception of the words as disyllabic, but increasing the duration of aspiration in "sport" did not; (2) perception of syllabicity appears to be categorical; and (3) there were no statistical differences among the four speech styles. [Work supported by NSF SOC75-10043.]

4:20

LL15. Stimulus range as a determinant of phoneme boundaries along synthetic consonant continua. Michael Studdert—Kennedy (Queens College, CUNY, Flushing, NY 11367 and Haskins Laboratories, 270 Crown St., New Haven, CT 06510)

Brady and Darwin (see Darwin, C.J. The Perception of Speech, in E.C. Carterette, and M.P. Friedman, *Handbook of Perception*, Academic, New York (in press), Vol. 7, report shifts in the phoneme boundary along a synthetic voicing continuum as a function of the range of stimuli presented within a test. The present study reports comparable shifts along synthetic voicing and place of articulation continua. Implications for the interpretation of adaptation boundary shifts are considered.

4:30

LL16. Perceptual adaptation to the duration of vowels preceding stop consonants. Paul D. Williams and Donald J. Sharf (Section of Speech and Hearing Sciences, The University of Michigan, 1111 E. Catherine St., Ann Arbor, MI 48104)

The affect of selective adaptation on the ability to identify stop consonants as being either voiced or voiceless solely on the basis of preceding vowel duration was tested. Stimuli consisted of computer-modified, real-speech sounds edited from a single utterance of /æd/ which varied in duration in 25-msec steps. Each subject classified recorded, random orderings of the stimuli as either AT or AD before and after periods of repetitive listening to long and short adapting stimuli. Calculated phonetic boundaries showed a mean shift toward the category of the adapting stimulus in both conditions, indicating a reduction in sensitivity for that category. The findings of independent, adaptable feature detectors for the voiced and voiceless categories and a greater resistance to adaptation for the voiced feature were comparable to results reported for adaptation studies employing voice onset time. [P.D. Eimas and J.D. Corbit, Cognitive Psychology 4, 99—109 (1973); P.D. Eimas *et al.*, Perception and Psychophysics 13, 247—253 (1973)].