27. Analysis of the Temporal Relationship Between Pitch Control and Articulatory Movements in the Realization of Japanese Word Accent by a Patient with Apraxia of Speech

Shinobu Masaki, Itaru F. Tatsumi,
and Sumiko Sasanuma

Many researchers have analyzed the speech abnormalities in patients with apraxia of speech (Deal & Darley, 1972; Johns & Darley, 1970; Kent & Rosenbek, 1982, 1983; LaPointe & Johns, 1975; Sasanuma, 1971; Shankweiler & Harris, 1966; Trost & Canter, 1974). Recently developed observational methods for articulatory movements, such as the fiberscope, the X-ray microbeam system, the labial-mandibular-movement transduction system using strain gauge, and electropalatography, have demonstrated incoordination in movements among various articulators (Fromm, Abbs, McNeil, & Rosenbek, 1982; Hardcastle, 1987; Itoh, Sasanuma, Hirose, Yoshioka, & Ushijima, 1980; Itoh, Sasanuma, & Ushijima, 1979). Incoordination similar to that found among different articulators may also be found between the movements of muscles used for pitch control on the one hand and those used for articulation on the other. Quantitative analyses of the temporal relationship between pitch control and articulatory movements may help to further elucidate the mechanism of speech abnormalities in apraxia of speech.

Japanese word accent is realized by the upward or downward transition of pitch between morae. Fujisaki (1977) analyzed the characteristics of the coordination between pitch control and articulatory movements for word accent in a normal speaker based on functional models of the generation

The authors are grateful to Mrs. Mariko Fujibayashi and Mr. Shoichi Tsuchida at Toranomon Hospital for their help with obtaining speech samples. Above all, we would like to thank our patient S.U. for her cooperation and forbearance.

process of pitch contour and formant trajectories. This method of analysis can be used to analyze the characteristics of the coordination between pitch control and articulatory movements in patients with apraxia of speech. The purpose of this study was to examine whether the temporal relationship between pitch control and articulatory movements was preserved in a patient with apraxia of speech.

METHOD

Subjects

The subjects were a 27-year-old female patient, S. U., with relatively pure apraxia of speech and five age- and sex-matched normal controls. All subjects were right handed and speakers of the Tokyo dialect.

S. U. suffered an embolic cerebrovascular accident resulting from heart disease 6 years before the data acquisition. A computerized tomography scan at 3 months post-onset showed circumscribed low-density areas in the cortex and subcortical white matter of the second and third frontal gyri, and the precentral gyrus of the left hemisphere.

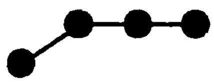
The *Standard Language Test of Aphasia* (Hasegawa, 1980), consisting of 26 subtests including Auditory Comprehension, Reading Comprehension, Oral Production, Writing, and Calculation, was administered to S. U. at 8 months post-onset. The results revealed that her scores were within the normal range in all modalities except Oral Production.

Speech abnormalities in S. U. were characterized by inconsistent segmental errors and disturbed control of duration for each segment. Segmental errors included consonant errors and vowel distortions. Disturbed control of duration occurred in the form of the lengthening of transitional or steady-state components and the temporal separation of syllables. No abnormal word accent pattern could be detected by simply listening to her speech.

Speech Materials

Speech materials were 16 three- to five-mora words beginning with [ai], all having a homogeneous accent pattern, that is, an upward pitch transition at the end of the initial mora in Tokyo dialect. In Table 27.1, examples of speech materials are shown with their pronunciation, accent pattern, orthographic representation in Japanese, and meaning. The accent pat-

TABLE 27.1. EXAMPLES OF SPEECH MATERIALS

Accent Pattern and Pronunciation	*Orthographic Representation in Japanese*	*Meaning*
 a- i- ma	合間	interval
 a- i- i- ɾo	藍色	indigo blue
 a- i- da- ga- ɾa	間柄	relation

tern is represented schematically by the binary (high/low) levels of subjective pitch corresponding to each mora.

The subjects were asked to repeat the model utterance of each word, which they heard presented on an audio tape recorder. A normal speaking rate of 5.0 mora/s and a slow speaking rate of 3.3 mora/s, which was the natural speaking rate of the patient S. U., were used. One utterance was recorded and analyzed for each word at each speaking rate.

Analyses

Onsets of accent and articulatory commands were estimated, and the temporal relationship between these commands was examined. Analyses were performed by a digital computer. The recorded materials were sampled at 10 kHz with an accuracy of 12 bits.

Estimation of Onset of Accent Command. Accent command was estimated by the analysis of pitch contour. Pitch contour was extracted from the speech wave by measuring intervals between peaks corresponding to the opening of the glottis.

Analysis of the pitch contour was based on the functional model of the generation process of pitch contour (Fujisaki, 1977). In this model, pitch contour was assumed to be composed of two components (Figure 27.1). One of these components was a voicing component, which was characterized by an initial rise and a gradual decay toward the end of an utterance. This component was represented by a response of a critically damped, second-order linear system that accepted an impulse. The other component was an accent component, which was characterized by a

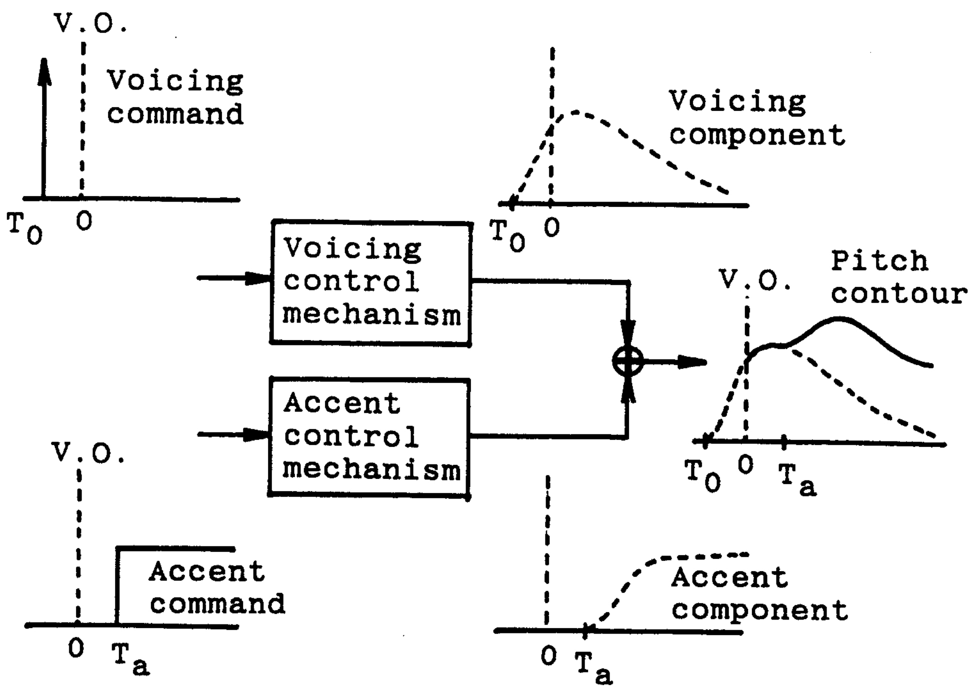


Figure 27.1. Functional model of generation process of pitch contour (Fujisaki, 1977). V.O., T_0 and T_a represent instants of voice onset, voicing command, and accent command, respectively.

smooth rise corresponding to upward transition of subjective pitch in order to realize Japanese word accent. This component was represented by a response of another critically damped, second-order linear system that accepted a stepwise transition of the subjective pitch.

Pitch contour was synthesized by the combination of these two components. Onset and magnitude of these commands were estimated in such a way that the pitch contour synthesized by the model gave the closest approximation to the observed contour.

Estimation of Onset of Articulatory Command. Onset of articulatory command for the second vowel [i] was estimated from the formant trajectories. The first three formant frequencies were extracted pitch-synchronously using the "Analysis-by-Synthesis" method (Fujisaki, Nakamura, & Yoshimune, 1970).

Estimation of articulatory command was based on the functional model of the coarticulatory process of vowel sequence (Fujisaki, 1977). The model describes the conversion process from a sequence of target formant frequencies for consecutive vowels into smoothed formant trajectories (Figure 27.2). This conversion process was represented by a response of a critically damped, second-order linear system that accepted a stepwise

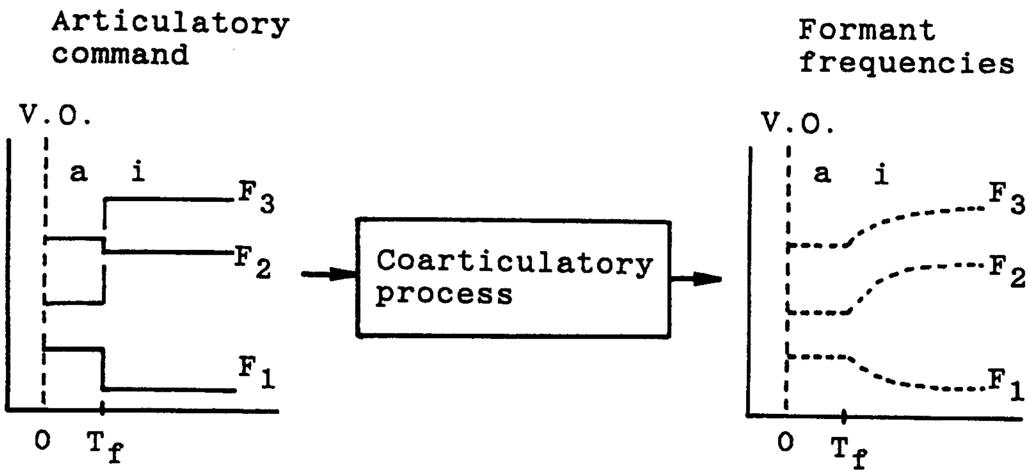


Figure 27.2. Functional model of coarticulatory process in the formant (F_1, F_2, and F_3) frequency domain for vowel sequence [ai] (Fujisaki, 1977). V.O. and T_f represent instants of voice onset and articulatory command, respectively.

transition of the target formant frequencies. The target formant frequencies and a common onset of their stepwise transitions were estimated as an articulatory command. The estimation of the articulatory command was performed in such a way that the formant trajectories synthesized by the model gave the closest approximation to the observed trajectories.

RESULTS

Figures 27.3 and 27.4 show examples of the estimated results of accent and articulatory commands for an utterance by a normal subject and the patient S. U., respectively. A Japanese word *aima,* meaning "interval," was uttered at a slow speaking rate. From top to bottom, speech wave, pitch contour, and formant trajectories are shown. The extracted pitch contour is represented by an array of dots. A broken line, which overlaps with the array of dots, represents the synthesized pitch contour that gives the closest approximation to the extracted pitch contour. The onset of the accent command is represented by a vertical line T_a. At the bottom, the first three formant trajectories are represented by arrays of dots, triangles, and squares. Broken lines represent the synthesized trajectories that give the closest approximation to the extracted formant trajectories. Stepwise solid lines represent estimated target formant frequencies. The onset of the articulatory command for the second vowel [i] is represented by a vertical line T_f.

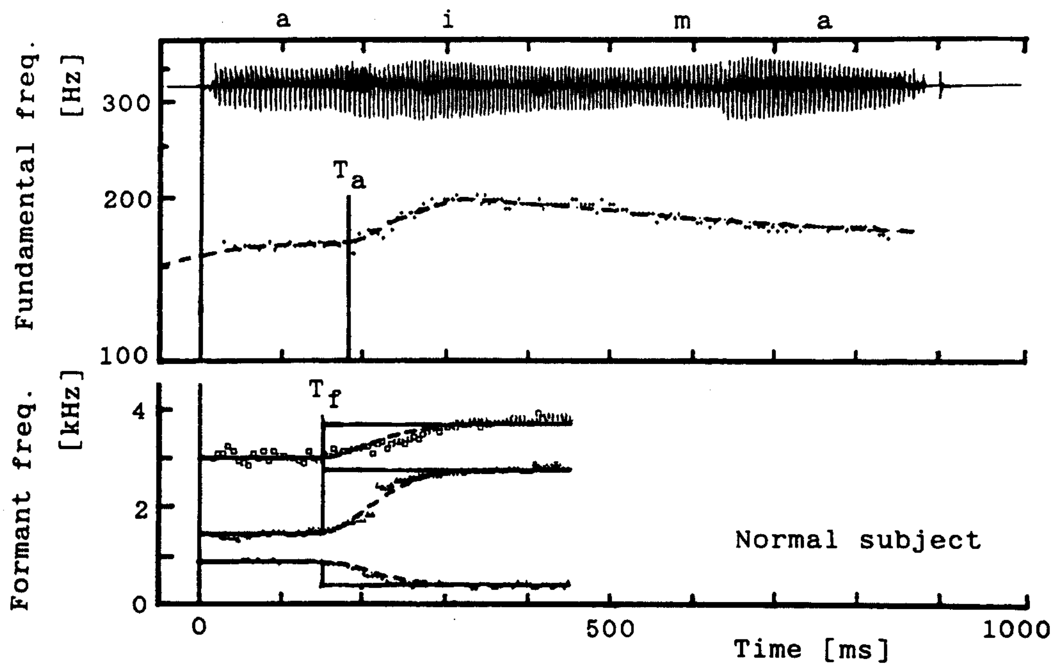


Figure 27.3. An example of extracted speech wave, pitch contour, and formant trajectories with estimated results of the accent command and articulatory command for a normal subject.

As shown in Figures 27.3 and 27.4, synthesized pitch contour and formant trajectories based on the models gave good approximations not only for the slow speaking rate but also for the normal speaking rate in all subjects including S. U. In these examples, an accent command T_a was delayed relative to an articulatory command T_f by 30 ms for the normal subject (Figure 27.3), whereas it was delayed by 130 ms for S. U. (Figure 27.4).

Figure 27.5 shows scatterplots of the temporal relationship between the accent command and the articulatory command. Abscissa T_f represents the duration from voice onset to onset of the articulatory command. Ordinate T_a represents the duration from voice onset to onset of the accent command. Circles represent the data points for the normal speaking rate at 5.0 mora/s, and triangles represent the data points for the slow speaking rate at 3.3 mora/s. In each speaking rate, filled marks represent pooled data points for the five normal subjects; open marks represent S. U.

For the normal speaking rate, accent commands in normal subjects occurred either before or after articulatory commands. The temporal relationship between these commands is specified by the value $T_a - T_f$. The range of the temporal relationship was between -60 ms and 70 ms as represented by slanted solid lines. For S. U., the temporal relationship was between -40 ms and 50 ms. All data points were within the normal range.

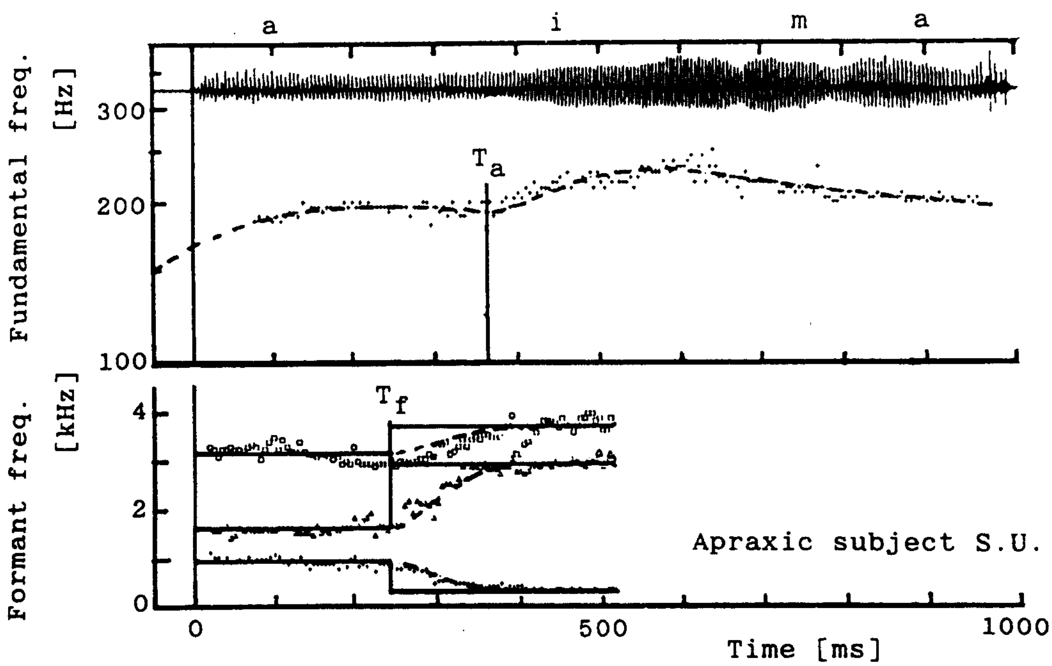


Figure 27.4. An example of extracted speech wave, pitch contour, and formant trajectories with estimated results of the accent command and articulatory command for the patient S. U.

As for the slow speaking rate, accent commands in normal subjects also occurred either before or after the articulatory commands. The range of the temporal relationship between these commands was between –50 ms and 80 ms. However, for S. U., all accent commands occurred after the articulatory commands. The range of the temporal relationship was specified as between 60 ms and 170 ms. In addition, the delay of the accent command relative to the articulatory command exceeded the normal range in 13 out of 16 instances.

DISCUSSION

The temporal relationship between pitch control and articulatory movements in our apraxic patient was within the normal range for the normal speaking rate. However, delay of pitch control relative to articulatory movements was observed for the slow speaking rate. The exact mechanism that caused the incoordination between pitch control and articulatory movements for the slow speaking rate but not the normal speaking rate is

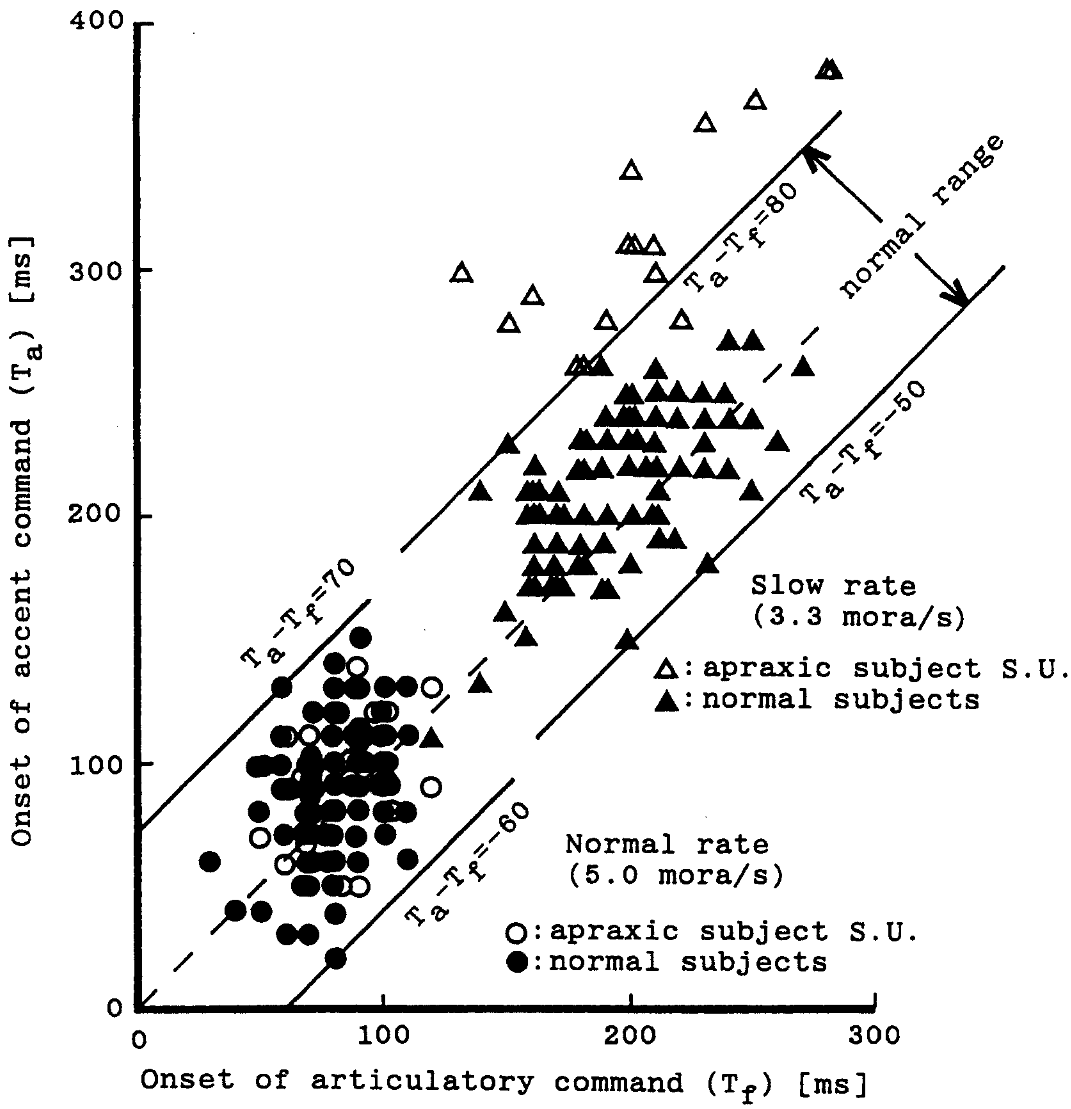


Figure 27.5. Scatterplot of the temporal relationship between the articulatory command (T_f) and accent command (T_a).

unclear. At the least, it appears that temporal coordination is influenced by speaking rate.

The existence of incoordination between pitch control and articulatory movements in the patient appears to be important. Until now, much evidence was reported indicating incoordination among multiple articulatory movements in patients with apraxia of speech. Furthermore, this incoordination was also observed between phonation control and articulatory movements by using measurement of voice onset time (Freeman, Sands, & Harris, 1978; Itoh et al., 1982; Kent & Rosenbek, 1983) and by using a combination of acoustic and electropalatographic analyses (Sugishita et al., 1987). In these reports, incoordination was observed in the form of delay of phonation control relative to articulatory movements. In the present study, incoordination as well as a tendency for glottal control delay relative to the articulatory movements was observed, suggesting that the results are an extension of these previous findings.

As for the prosodic features of patients with apraxia of speech, Darley, Aronson, and Brown (1975) mentioned: "Some prosodic alterations—pauses, slowed rate, equalization of stress—may also appear, probably in compensation for the continuing articulatory difficulty" (p. 262). However, the nature of the pitch contour, which is one of the important components of prosody, has not been investigated quantitatively, chiefly because methodology has not been established. The results of this study suggest that pitch contour abnormality is caused by the same mechanism as articulation or the segmental feature abnormality—that is, by the incoordination of muscle control. These results in turn indicate the possibility that at least some components of abnormal prosody can be considered to be a "primary" rather than a "compensatory" or "secondary" phenomenon.

REFERENCES

Darley, F. L., Aronson, A. E., & Brown, J. R. (1975). *Motor speech disorders.* Philadelphia: W. B. Saunders.

Deal, J. L., & Darley, F. L. (1972). The influence of linguistic and situational variables on phonemic accuracy in apraxia of speech. *Journal of Speech and Hearing Research, 15,* 639-653.

Freeman, F. J., Sands, E. S., & Harris, K. S. (1978). Temporal coordination of phonation and articulation in a case of verbal apraxia: A voice onset time study. *Brain and Language, 6,* 196-211.

Fromm, D., Abbs, J. H., McNeil, M. R., & Rosenbek, J. C. (1982). Simultaneous perceptual-physiological method for studying apraxia of speech. In R. H. Brookshire (Ed.), *Clinical aphasiology* (Vol. 12, pp. 251-262). Minneapolis, MN: BRK Publishers.

Fujisaki, H. (1977). Functional model of articulatory and phonatory dynamics: Temporal organization of articulatory and phonatory controls in realization of Japanese word accent. In M. Sawashima & F. S. Cooper (Eds.), *Dynamic aspects of speech production* (pp. 27-34). Tokyo: University of Tokyo Press.

Fujisaki, H., Nakamura, N., & Yoshimune, K. (1970). Analysis, nominalization and recognition of sustained Japanese vowels. *Journal of the Acoustical Society of Japan, 26,* 152-154.

Hardcastle, W. J. (1987). Electropalatographic study of articulation disorders in verbal dyspraxia. In J. H. Ryalls (Ed.), *Phonetic approaches to speech production in aphasia and related disorders.* San Diego, CA: College-Hill Press.

Hasegawa, T. (Ed.). (1980). *Aphasia: Mechanisms and management.* Tokyo: Kongo Shuppan. (in Japanese)

Itoh, M., Sasanuma, S., Hirose, H., Yoshioka, H., & Ushijima, T. (1980). Abnormal articulatory dynamics in a patient with apraxia of speech: X-ray microbeam observation. *Brain and Language, 11,* 66–75.

Itoh, M., Sasanuma, S., Tatsumi, I. F., Murakami, S., Fukusako, Y., & Suzuki, T. (1982). Voice onset time characteristics in apraxia of speech. *Brain and Language, 17,* 193–210.

Itoh, M., Sasanuma, S., & Ushijima, T. (1979). Velar movements during speech in a patient with apraxia of speech. *Brain and Language, 7,* 227–239.

Johns, D. F., & Darley, F. L. (1970). Phonemic variability in apraxia of speech. *Journal of Speech and Hearing Research, 13,* 556–583.

Kent, R. D., & Rosenbek, J. C. (1982). Prosodic disturbance and neurologic lesion. *Brain and Language, 15,* 259–291.

Kent, R. D., & Rosenbek J. C. (1983). Acoustic patterns of apraxia of speech. *Journal of Speech and Hearing Research, 26,* 231–249.

LaPointe, L. L., & Johns, D. F. (1975). Some phonemic characteristics in apraxia of speech. *Journal of Communication Disorders, 8,* 259–269.

Sasanuma, S. (1971). Speech characteristics of a patient with apraxia of speech: A preliminary case report. *Annual Bulletin of Research Institute of Logopedics and Phoniatrics, University of Tokyo, 5,* 85–89.

Shankweiler, D., & Harris, K. S. (1966). An experimental approach to the problems of articulation in aphasia. *Cortex, 2,* 277–292.

Sugishita, M., Konno, K., Kabe, S., Yunoki, K., Togashi, O., & Kawamura, M. (1987). Electropalatographic analysis of apraxia of speech in a left hander and in a right hander. *Brain, 110,* 1393–1417.

Trost, J. E., & Canter, G. J. (1974). Apraxia of speech in patients with Broca's aphasia: A study of phoneme production accuracy and error patterns. *Brain and Language, 1,* 63–79.