Ann. Bull. RILP, No. 13, 123-132 (1979)

VOICE ONSET TIME CHARACTERISTICS OF APRAXIA OF SPEECH

Motonobu Itoh*, Sumiko Sasanuma**,
Itaru F. Tatsumi * and Yo Kobayashi*

Apraxia of speech has been defined as a disorder of motor programming for speech (Darley, Aronson and Brown, 1975). In order to obtain an accurate picture of this syndrome, we undertook an in-depth analysis of the movements of the articulators during speech in a patient diagnosed as having apraxia of speech using a fiberoptic system and a computer-controlled X-ray microbeam system (Itoh, Sasanuma, Hirose, Yoshioka and Ushijima, 1978; Itoh, Sasanuma and Ushijima, 1979). Results clearly indicated that there was a defective temporal organization of articulation in the apraxic patient. Based on these results, we predicted certain alterations in the timing of glottal activity relative to supraglottal articulatory adjustments in apraxia of speech. With this expectation, we examined the timing control of the laryngeal and supra-laryngeal articulatory adjustments in apraxic patients by measuring voice onset time (VOT).

1. Method

1.1. Subject

We obtained the VOT data from three groups of subjects. They were four brain-damaged adult subjects with apraxia of speech, four young adult subjects and five aged subjects. The apraxic subjects were patients of the Tokyo Metropolitan Geriatric Hospital and were between the ages of 40 and 61 years. All of them exhibited a conspicuous impairment in articulation as well as in prosody due to apraxia of speech. Two of these patients had mild aphasia in addition to apraxia of speech. According to the recent grouping of aphasia by Schuell (Jenkins, Jimenez-Pabon, Shaw and Sefer, 1975), these two patients can be classified as aphasia with dysfluency. Table 1 summarizes the age, cause of brain damage, site of lesion and post onset time for the four apraxic patients. The young adult subjects were staff members of the Tokyo Metropolitan Institute of Gerontology and their age ranged from 25 to 32 years. The aged subjects were volunteers from local senior-citizens groups and their age ranged from 66 to 72 years. All subjects were male and, with the exception of the apraxic subjects, had normal speech and hearing for their age.

---

* Tokyo Metropolitan Institute of Gerontology
** Yokohama National University/ Tokyo Metropolitan Institute of Gerontology.

Table 1. The age, cause of the brain damage, site of lesion and post onset time for the four apraxic patients.

Subject	Age	Cause of Brain Damage	Site of Lesion**	Post Onset Time
C1	61	CVA	Near the anterior tip of the Sylvian fissure of the left hemisphere, both cortical and subcortical.	6 years 7 months
C2	66	CVA	Near the inferior frontal gyrus of the left hemisphere, both cortical and subcortical.	4 years 11 months
C3*	45	CVA	Near the anterior tip of the Sylvian fissure of the left hemisphere, both cortical and subcortical.	6 months
C4*	40	Brain tumor	The parietal lobe of the left hemisphere.	2 months

* With mild aphasia

** Identified by CT scan

1.2. Procedure

Each subject read a randomized list of the Japanese monosyllables /de/, /te/, /ge/, and /ke/. The list contained 25 occurrences of each syllable for a total of 100 items in the test. Subjects were instructed to indicate when they misread and noticed the error. If any subject indicated such an error, he was asked to re-read the syllables which he had misread.

Wide-band spectrograms of the 25 productions of each syllable were made from the recordings. The VOT measurements were made according to the procedure used by Lisker and Abramson (1964). That is, the VOT was measured as the interval between the first vertical striation representing the vocal fold vibration and the onset of energy "burst" representing the release of an articulatory constriction (Figure 1). When the vocal fold vibration precedes the release as seen in /de/, the VOT is given a negative value and is called voicing lead. On the other hand, when the release precedes the vocal fold vibration as in /te/, the VOT value is positive and is called voicing lag. Identification of the starting point of the vertical striation and the energy "burst" was made by two of the investigators. Agreement was achieved on 96% of the identification tasks. Utterances on which agreement was not achieved were discarded. VOT measurements obtained from the spectrograms of each syllable were grouped into 10 msec intervals and frequency distributions were plotted for every subject.

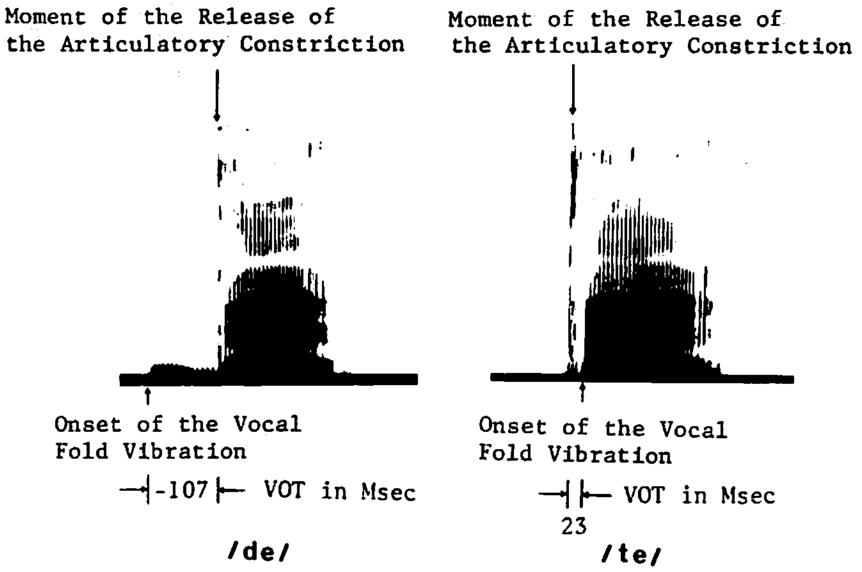

Fig. 1. Examples of VOT measurements.

2. Results

2.1. Group comparison

Figure 2 presents the pooled data for each of the three subject groups in the form of frequency distributions of the VOT values for /de/ and /te/ (left), and /ge/ and /ke/ (right). The abscissa represents the VOT values in msec and the ordinate represents the total number of responses. The black columns indicate the VOT values for the voiced targets /de/ and /ge/, and the blank columns indicate those of the voiceless targets /te/ and /ke/. The descriptive statistics for the VOT distributions associated with productions of /de/, /ge/, /te/ and /ke/ for each subject group are included in Table 2 (because of the small number of subjects in each group, statistical analysis was not conducted).

As Figure 2 and Table 2 indicate, the normal young and aged subjects' voiced and voiceless productions are distributed into relatively discrete areas of VOT although there is a sporadic occurrence of overlap between the VOT values for the two phonetic categories in the distribution of the aged group. That is, the majority of the VOT values for the voiced targets /de/ and /ge/ lie in the voicing lead category and all of the VOT values for the voiceless targets /te/ and /ke/ occur in the voicing lag category. It is also evident that there is a reliable difference in the size of range between the two voicing categories, with a greater dispersion of production along the VOT continuum for the voiced stops in comparison with the voiceless stops. Furthermore, the VOT values of the voicing lead associated with production of the voiced velar stop /g/ tended to be smaller than those of the voiced alveolar stop /d/. On the other hand, the VOT values of the voicing lag associated with production of voiceless velar stop /k/ tended to be greater than those of the voiceless alveolar stop /t/.

Contrary to this, the group of apraxic subjects exhibited a considerable overlap in VOT distribution between the voiced and voiceless cognates. It is also clear that there is a greater dispersion of productions along the continuum for both the voiceless and voiced stops in comparison with the other subject groups. However, some patterns observed for the normal subject groups were also found for the apraxic group. One of the patterns noted was that the range of the VOT distribution of the voiced stop was larger than that of the voiceless stop. Another pattern observed for the normal subject group and also found in the apraxic group was that the VOT values of the velar stops systematically differed from those of the alveolar stops.

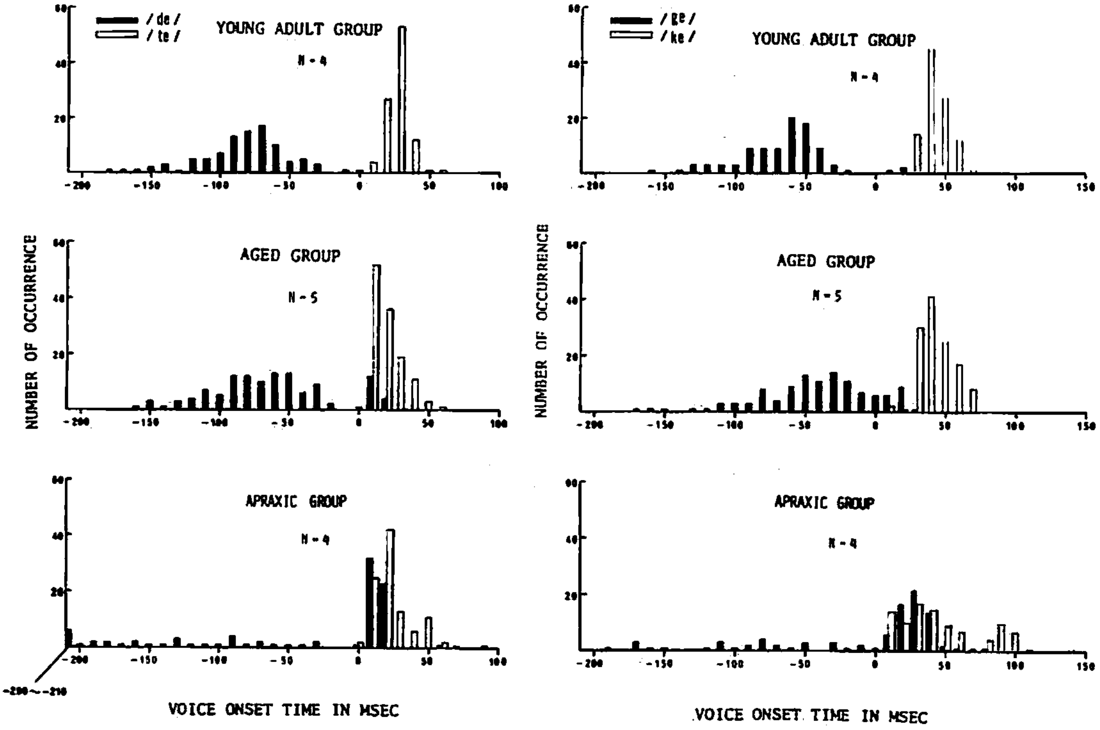

Fig. 2. VOT distributions of three subject groups for /de/ and /te/ (left), and /ge/ and /ke/ (right).

Table 2. Mean, mode, minimum, maximum and range of VOT distributions for three subject groups.

Subject Group	Syllable	Mean	Mode	Minimum	Maximum	Range
Young Adult N = 4	de	-82.24	-70	-180	0	180
	ge	-66.02	-60	-160	20	180
	te	27.26	30	10	60	50
	ke	44.45	40	30	70	40
Aged N = 5	de	-63.46	-55	-160	20	180
	ge	-40.70	-30	-170	30	200
	te	20.54	10	10	60	50
	ke	43.74	40	10	70	60
Apraxic N = 4	de	-47.51	10	-290	70	360
	ge	-7.16	30	-190	80	270
	te	25.00	20	0	90	90
	ke	47.13	30	10	140	130

(In msec)

2.2. Individual comparison

Figures 3, 4, 5 and 6 present individual data of each subject from the three subject groups. Figure 3 gives the VOT production distributions for /de/ and /te/ (left), and /ge/ and /ke/(right) for each of the four young adult subjects. As this figure indicates, each subject's voiced and voiceless productions are distributed into two discrete areas of VOT with few exceptions. There is no overlap of the VOT values between the two phonetic categories. Although Subject A1's distribution pattern differs somewhat from those of the other three subjects, the discreteness of the temporal categories for the voiced and voiceless cognates sharing the same place of articulation was clearly maintained in this subject.

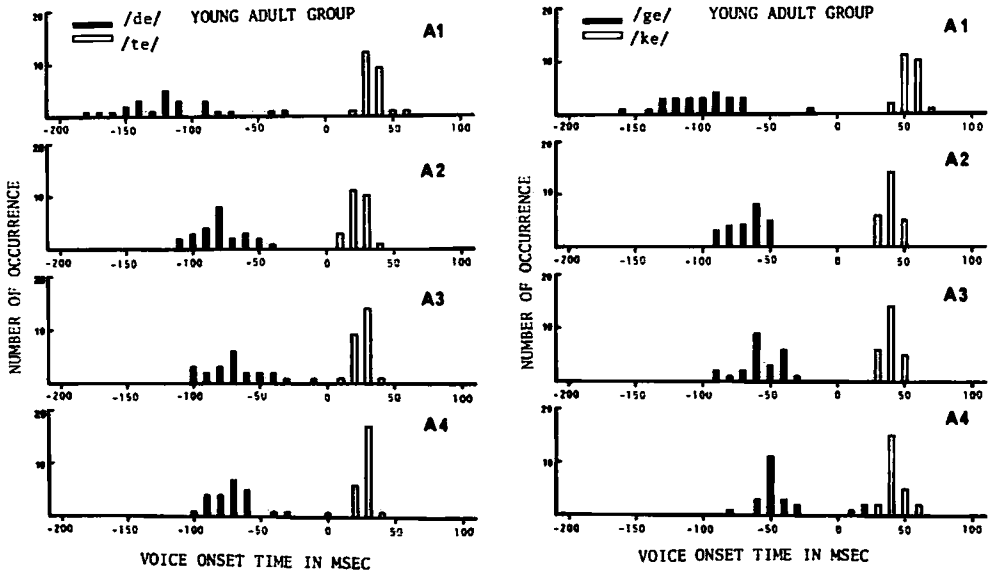

Fig. 3. Individual VOT distributions of the young adult group for /de/ and /te/(left), and /ge/ and /ke/ (right).

Figure 4 presents the distribution patterns of the VOT values for each of the five aged subjects. The distributions of the VOT values for the voiced and voiceless cognates are relatively discrete with a sporadic overlap. It can be seen that the discreteness of the distributions for the two phonetic categories is obscured to some degree in the productions of B2 and there is a minor overlap of the VOT values between /de/ and /te/ in the productions of Subjects B1 and B5.

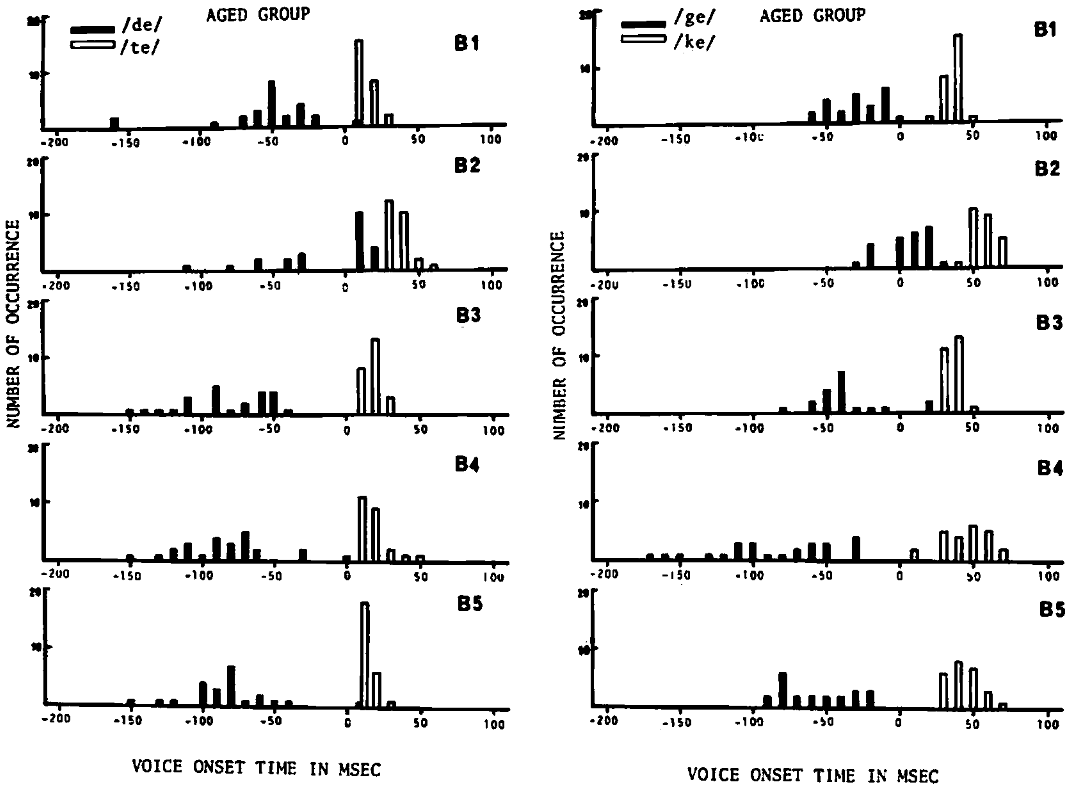

Fig. 4. Individual VOT distributions of the aged group for /de/ and /te/ (left), and /ge/ and /ke/ (right).

Figures 5 and 6 present distribution patterns of the VOT values for each of the four apraxic subjects for /de/ and /te/, and /ge/ and /ke/ respectively. The VOT productions of Subjects C1, C2 and C3 exhibited a considerable overlap between the two phonetic categories. In addition, C1 did not show any voicing lead, though this was observed for /de/ and /ge/ in the normal subjects. Furthermore, the range of C2's distribution, especially for /ge/ and /ke/, is very narrow and there is a complete overlap of distribution between the two phonetic categories. Although Subject C4 showed only a slight overlap of VOT distributions between /de/ and /te/, his productions of the voiced targets, especially /de/, were characterized by an abnormally long voicing lead.

3. Discussion

The results obtained from the normal subjects seem to indicate that the timing relation between the articulatory release and the onset of the vocal fold vibration for the voiced and voiceless cognates is distinctively controlled in the normal subjects. Although there has been no systematic study evaluating VOTs in Japanese, VOTs have been extensively studied in normal speakers in different languages (e. g., Lisker and Abramson, 1964). These studies showed that the VOT values for voiced and voiceless stops are distributed into two discrete areas, which is

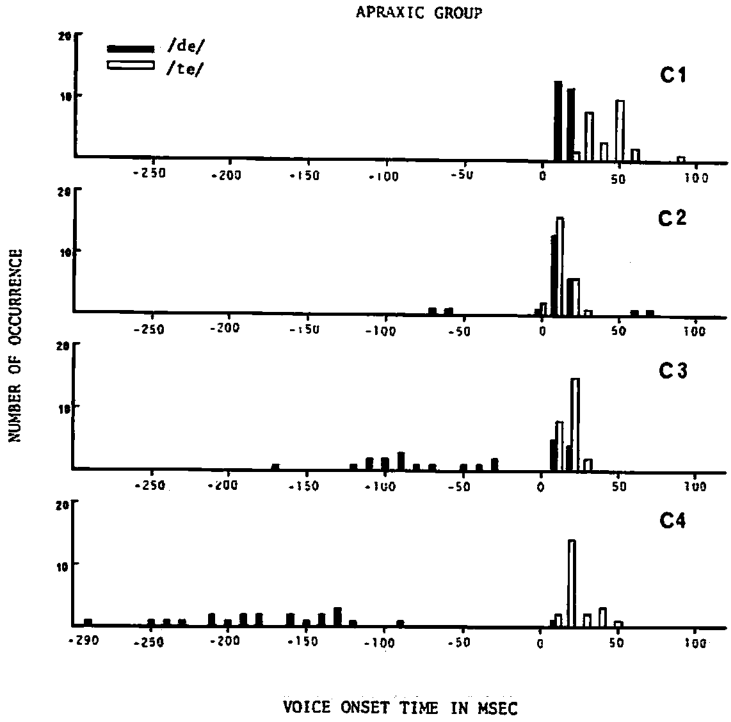

Fig. 5. Individual VOT distributions of the apraxic group for /de/ and /te/.

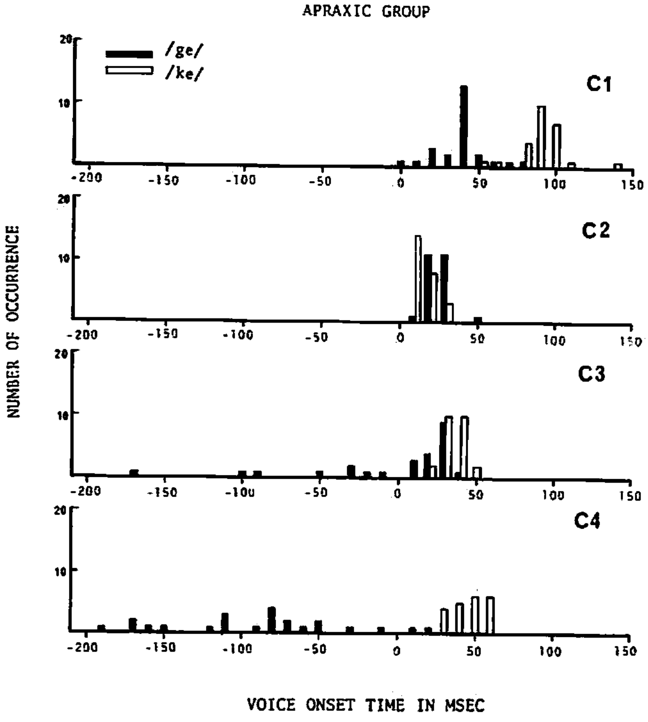

Fig. 6. Individual VOT distributions of the apraxic group for /ge/ and /ke/.

consistent with the results of the present study. The finding of the larger values of voicing lag for the velar voiceless stop /k/ than for the alveolar voiceless stop /t/ and the smaller values of voicing lead for /g/ than /d/ in the present study is also in accord with the results reported in the above studies.

In contrast to the behavior of the normal subjects, the apraxic subjects showed pathological distributions of VOT values. This result can be interpreted to indicate that the control over the timing of laryngeal and supra-laryngeal articulatory events is disturbed in this subject group. There has been only one study evaluating VOT in the speech of a patient with apraxia of speech. Freeman, Sands and Harris (1978) measured VOT for syllable initial stops in isolated words in an apraxic patient. They found that 1) the VOTs of the apraxic patient differed markedly from those of normal subjects, 2) the apraxic productions did not include voicing lead for voiced stops, 3) lag times for voiced stops were longer than normal, while those for voiceless stops were shorter than normal, and, as a result, 4) there is a compression of the two categories and a marked overlap. Freeman et al.point out that the VOT distributions produced by their apraxic subject were quite similar to those of children in the early developmental stages, as reported by Kewley-Port and Preston (1974). Based on these observations, Freeman et al. suggest that the articulation of the apraxic subject can be considered as a sort of regression to a primitive stage. However, they caution that it remained to be determined whether the problem of their apraxic subject was an idiosyncratic or common feature of apraxia of speech. The VOT distribution patterns of C2 in the present study are quite similar to those of the apraxic subject in Freeman et al.'s study. On the other hand, C1's lag time for the voiceless stop was longer than the normals', which is in sharp contrast with the reslut of Freeman et al.'s subject. Furthermore, the distribution patterns of C3 and C4, who had mild aphasia in addition to apraxia of speech, differed markedly from those of Freeman et al.'s subject. In view of these observations, it appears that the characteristics of distributions of Freeman et al.'s subject does not represent a common feature of apraxic defects.

Blumstein, Cooper, Zurif and Caramazza (1977) reported that all of their anterior aphasics showed an overlap between the voiced and voiceless categories along the VOT continuum in the productions of syllable initial stops in monosyllable words. Based on this result, they suggest that this phenomenon observed in the anterior aphasics seems to reflect a defect in the articulatory programming of speech sounds. We believe that Blumstein et al.'s subjects had a deficit of articulation similar to our apraxic subjects.

4. Final comments

The findings of our previous studies examining the movements of the articulators during speech in an apraxic patient allowed us to conclude that inconsistent articulatory errors, which were the major characteristics of apraxic speech, could be attributable to an impairment of time programming for the appropriate phoneme rather than to a defect in selecting an appropriate phoneme. The present results seem to offer additional support for such a view. However, further observation of more patients is certainly in order since only a small number of apraxic subjects were examined in the present study and they showed rather great individual differences as well. Furthermore, it is also necessary to compare the VOT distributions of apraxic patients with those of fluent aphasics in order to elucidate the nature and underlying mechanisms of nonfluent speech in apraxia of speech.

Acknowledgment

This study was supported in part by Grants in Aid for Scientific Research (No. 310707), Japanese Ministry of Education.

References

Blumstein, S. E., W. E. Cooper, E. B. Zurif, and A. Caramazza, 1977. The perception and production of voice-onset time in aphasia. Neuropsychologia, 15, 371-383.

Darley, F. G., A. E. Aronson, and J. R. Brown, 1975. Motor Speech Disorders. W. B. Saunders.

Freeman, F. J., E. S. Sands, and K. S. Harris, 1978. Temporal coordination of phonation and articulation in a case of verbal apraxia: A voice onset time study. Brain and Language, 6, 106-111.

Itoh, M., S. Sasanuma, H. Hirose, H. Yoshioka and T. Ushijima, 1978. Articulatory dynamics in a patient with apraxia of speech: X-ray microbeam observation. Annual Bulletin, Research Institute of Logopedics and Phoniatrics, University of Tokyo, No. 12, 87-96.

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

Jenkins, J., E. Jimenez-Pabon, R. E. Shaw, and J. W. Sefer, 1975. Schuell's Aphasia in Adults: Diagnosis, Prognosis, and Treatment. Harper and Row.

Kewley-Port, D. and M. S. Preston , 1974. Early apical stop production: A voice onset time analysis. Journal of Phonetics, 2, 195-210.

Lisker, L., and A. S. Abramson, 1964. A cross-language study of voicing in initial stops: Acoustical measurements. Word, 20, 384-422.