Journal of Speech and Hearing Research, GERRATT, *Volume* 26, 297-304, June 1983

Research Note

FORMANT FREQUENCY FLUCTUATION AS AN INDEX OF MOTOR STEADINESS IN THE VOCAL TRACT

BRUCE R. GERRATT
VA Medical Center, West Los Angeles, and UCLA School of Medicine

Involuntary movement of the articulatory structures can interfere with the accurate placement of the articulators during consonant production and may also result in distortion of vowel quality. An acoustic method was used to assess motor steadiness in the vocal tract musculature superior to the glottis during vowel production by five subjects with abnormal involuntary orofacial movements associated with tardive dyskinesia and 10 normal subjects. A linear predictive coding technique of spectral analysis yielded formant frequencies from the sustained productions of /ɑ/. Based on the premise that changes in vocal tract configuration can be measured as changes in formant frequency, the sequential segment-to-segment fluctuations of the second formant frequency of these vowel samples were computed and used as an index of motor steadiness. Results showed that formant frequency fluctuation measures for four of the five tardive dyskinetic patients were substantially larger than those of the normal subjects, indicating a reduction of motor steadiness in these four subjects. Factors influencing the validity of this procedure and implications for its use are discussed.

The term *dysarthria* implies a lack of motor control of the vocal tract musculature during speech production. Measurement of motor steadiness in the vocal tract is of particular importance in patients who have difficulty maintaining the postural stability of the structures necessary for adequate speech production. Involuntary movement of the articulatory structures during the steady-state portion of a vowel can result in a distortion of vowel quality. Abnormal movements also may interfere with the accurate placement of the articulators during consonant production, resulting in imprecise consonants.

These two articulatory difficulties can combine to reduce the speech intelligibility seriously. For example, Darley, Aronson, and Brown (1969a, 1969b) identified a number of deviant perceptual dimensions in the disordered speech of 30 patients having dystonia, a hyperkinetic disorder caused by damage in the extrapyramidal system. In a ranking by severity of deviation, three of the

0022-4685/83/2602-0297$01.00/0

four most prominent deviant dimensions were related to articulatory inefficiency: imprecise consonants, distorted vowels, and irregular articulatory breakdown. The first two of these dimensions were also highly correlated with judgments of intelligibility. Apparently, involuntary movements affecting the orofacial structures interfered with accurate movement of the articulators and contributed to the breakdown of speech production. Quantification of the reduction of speech motor control in dysarthric patients is important to define objectively the deviant speech characteristics as well as to provide indices for the assessment of therapeutic intervention.

This paper presents information on a noninvasive method to measure the degree of motor steadiness in the vocal-tract musculature superior to the glottis, based on the premise that a change in vocal tract configuration can be measured as a change in formant frequency. To the extent that the supraglottal vocal tract configuration remains constant during the production of a vowel, little sequential segment-to-segment change in formant frequency values is expected to occur. The measurement of formant frequency fluctuation is an overall assessment of sequential segment-to-segment variation in formant frequency and as such is an index of the degree of the moment-to-moment change or fluctuation in vocal tract configuration.

In this study, formant frequency fluctuation was measured from the recorded acoustic waveforms of normal speakers and speakers having abnormal involuntary orofacial movements associated with tardive dyskinesia, a hyperkinetic disorder resulting from prolonged exposure to phenothiazines and related antipsychotic drugs. Orofacial dyskinesia is the most familiar characteristic of tardive dyskinesia (Marsden, Tarsy, & Baldessarini, 1975), and the symptoms of this disease may include dysarthria (Maxwell, Massengill, & Nashold, 1970; Portnoy, 1979; Gerratt, Fisher, & Goetz, Note 1). The presence of involuntary orofacial movements in these patients made it reasonable to expect that the formant patterns obtained from their acoustic spectra might differ from those obtained from normal speakers. The major purpose of the study was to assess this method of examining the motor steadiness of the supraglottal vocal tract musculature by measuring the formant frequency fluctuation in vowels produced by speakers having hyperkinetic dysarthria.

METHOD

Subjects

The formant frequency fluctuation data were obtained from productions of the sustained vowel /ɑ/ by five tardive dyskinetic speakers and 10 normal speakers. The four women and one man who comprised the pathologic group had a median age of 51:6 (yrs:mos) (range 27:2–72:1), and the normal group, five women and five men, had a median age of 52:0 (range 26:3–74:8).

A neurologist diagnosed all five pathologic speakers as having tardive dyskinesia with general trunk and limb chorea as well as abnormal involuntary movement of the lips, face, and jaw. Connected speech samples from these five subjects were evaluated by two speech pathologists who were familiar with dysarthric speech. All subjects were found to have a motor speech disorder consistent with hyperkinetic dysarthria as defined by Darley et al. (1969a). At the time of testing, these patients were receiving antipsychotic drugs for psychiatric illnesses; however, the doses of these drugs had not been altered for at least 6 weeks prior to the speech production task used in this study.

Procedure

The production of speech tasks was carried out in a sound-treated room, and all speech was recorded on an audiotape recorder (Ampex 600B). Each subject was seated and a head-worn microphone (Shure SM12A) was positioned a constant distance of 20 cm from the lips.

Each subject was asked to take a deep breath and to sustain the vowel /ɑ/ at a conversational level of loudness for as long, steadily, and clearly as possible. In addition, each subject was instructed to maintain the same vowel quality throughout the entire production. This procedure was demonstrated by the investigator, and the subject was given practice trials until he or she showed understanding of the task. Following this, three productions were recorded, and a segment 3 sec in duration from the middle portion of the longest production was used for subsequent analysis.

Acoustic Analysis

The measurement accuracy of the FFF procedure obviously depends on the accuracy of the method used to determine the formant frequency values. The analysis of each sample in this study was performed using a linear prediction technique (LPC), one of the most commonly used methods of speech analysis since it was first introduced by Atal and Schroeder (1967). According to Schroeder (1982), the success of this technique can be ascribed to two facts: Most speech sound characteristics can be extracted from the vocal tract modelled as an all-pole structure, and these characteristics can be derived quickly at reasonable cost using digital systems. For the purposes of this study, another important feature is that LPC analysis provides the fine temporal resolution necessary for measuring rapidly varying formant frequencies that cannot be obtained from spectrograms.

The formant frequencies were calculated using the autocorrelation method of linear prediction based on the work of Markel and Gray (1976). Basically, this method analyzes the speech wave by predicting the speech sample as a linear combination of a number of previous samples. Two major considerations in the use of this analysis were the choices of the number of prediction coefficients (filter length) and the length of the analysis window. Markel (1972) suggested that a filter length of 14–16 coefficients is a logical choice when sampling at 10 kHz.

If too many filter coefficients are computed, spurious peaks that do not represent formants may appear, whereas if too few coefficients are used, true formants many be missed. Empirical observation in this study revealed a filter length of 14 coefficients to provide clearly defined formant peaks. The choice of analysis window length also affects the ability to resolve formant peaks. An analysis that uses a window too short in duration may miss closely spaced formants, while the use of an excessively long window may miss strong formant peaks because of frequency averaging over the time interval of the window (Markel, 1972).

For this analysis, 256 waveform points per frame, representing 25.6 msec of the acoustic signal, were selected and a Hamming window was applied. Since FFF is designed to measure the variability of formant frequencies, a frame rate of 12.8 msec was used to increase the data per unit time and thereby minimize the omission of formant change during quick articulatory movements. Thus, after analysis of the first 25.6-msec segment, analysis of a second 25.6-msec segment was performed, commencing 12.8 msec after the beginning of the first segment. Therefore, each segment to be analyzed consisted of the last half of the preceding waveform segment and the initial half of the following segment.

Finally, formant frequency values from each segment were determined with the use of an algorithm for picking the peaks of the smoothed spectra (Markel & Gray, 1976). According to Wakita (1976), peak picking of the smooth spectral envelope will give correct results approximately 90% of the time. In this study, the spectral analysis procedure occasionally failed to pick peaks in the spectrum. In addition, a peak was rejected as spurious if it was found to have a bandwidth greater than 300 Hz. In order to reduce the likelihood of such missing values for any of the 25.6-msec segments, the analysis of the segments was performed every 12.8 msec until 78 consecutive segments without missing values were obtained, representing approximately 1 sec of the digitized vowel sample. For three tardive dyskinetic and two normal subjects, it was not possible to obtain values for all 78 segments. In these cases, no more than six segments were missed, and the missed segments usually occurred at the end and/or the beginning of the 1-sec sample.

As an example, Figure 1 presents the output derived from the analysis of 10 consecutive waveform segments. Although the analysis yielded values for many peaks in the spectrum, only the peak representing the second formant was used in the computation of formant fluctuation measures. The choice of F_2 was somewhat arbitrary, since the entire supraglottal vocal tract has some influence on all the formants (Fant, 1973). However, Peterson and Barney's (1952) data for F_1, F_2, *and* F_3 of /i, ɪ, ɛ, æ, ɑ, ɔ, u, ʊ, ʌ, ɝ/ produced by adult males showed a greater range and greater standard deviations for F_2 across all 10 vowels than for either F_1 or F_3. Thus, the greater variability of F_2 appears to reflect a greater sensitivity of this formant to changes in the shape of the vocal cavities, at least for the production of these vowels.

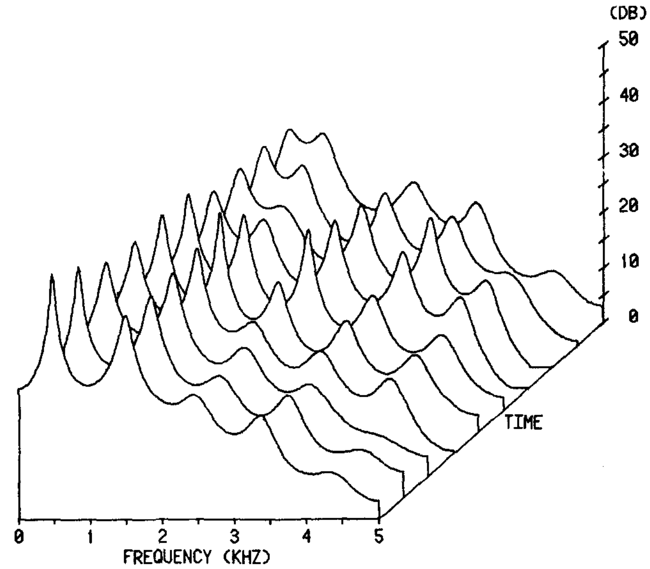

FIGURE 1. An example of spectral analysis of 10 consecutive waveform segments from the vowel /ɑ/. The time interval between each segment is 12.8 msec.

After performing the analyses, the F_2 frequency values were used to compute four measures: (a) mean F_2, (b) standard deviation of F_2, (c) mean F_2 formant frequency fluctuation ($\bar{x}$FFF), and (d) F_2 formant frequency fluctuation ratio (FFFR). Mean FFF was obtained by calculating the absolute magnitude of difference between the second formant frequencies for each consecutive pair of 25.6-msec segments and computing the average of these differences. FFFR was computed using the following procedure:

$$\text{FFFR} = \frac{\bar{x}\text{FFF}}{\bar{x}\text{FF}} \times 1000$$

where $\bar{x}$FFF is mean formant frequency fluctuation and $\bar{x}$FF is the mean formant frequency. These computational procedures for $\bar{x}$FFF and FFFR are conceptually similar to those employed in the study of vocal jitter (cf. Horii, 1979).

Validity Testing

It is obvious that any measurement technique involves some degree of inaccuracy. Recently, Monsen and Engebretson (1983) reported that the accuracy of linear prediction analysis is approximately ±60 Hz for the first three formants of 90 synthetic vowels representing a variety of different problems encountered in the measurement of formant frequency. Consequently, it was important to determine the accuracy of the particular methods used in the present study.

Vowels having constant second formant frequencies. The validity of the method of analysis in these measurements was first tested using vowels having constant formant frequencies to determine how much formant frequency fluctuation may occur as a result of error in the system. Four vowels, each 1 sec in duration, were synthesized using a digital speech synthesizer similar to that reported by Klatt (1980) for four conditions formed by a factorial of sex by fundamental frequency:

1. the vowel /ɑ/ of a female speaker with a constant F_0
2. the vowel /ɑ/ of a female speaker with a varying F_0
3. the vowel /ɑ/ of a male speaker with a constant F_0
4. the vowel /ɑ/ of a male speaker with a varying F_0.

Since most of the pathologic speakers manifested some degree of variability in fundamental frequency, an exaggerated amount of variability in F_0 was used to examine its influence on the accuracy of the formant frequency fluctuation measures. Specifically, F_0 for two of the vowels varied over 400 msec by 10 Hz above and below the mean F_0. The F_1, F_2, and F_3 values for the female speaker were 820, 1345, and 3005 Hz, respectively, while the mean F_0 was 220 Hz. The formant values for the male speaker were 725, 1229, and 2509 Hz, and the mean F_0 was 110 Hz. All four vowels were produced with unvarying formant frequencies for their entire duration. Thus, any formant frequency fluctuation which was measured subsequently can be assumed to have occurred as a result of system error. These test vowels were recorded, filtered, and digitized, and sequential spectral analysis was performed in a manner identical to that reported for the analysis of real speech.

Vowels having varying second formant frequencies. To determine the accuracy of this analysis in measuring change in F_2, eight additional synthesized vowels were analyzed. The acoustic parameters of the four vowels used in the validity testing described above were used; however, this time the additional factor of F_2 variation was included. Four of the eight vowels had F_2 frequencies which varied around the mean F_2 by 400 Hz, while the F_2 frequencies of the other four vowels varied by 800 Hz, as presented in Figure 2. Synthesized vowels thus were produced for the eight conditions formed by a 2 × 2 × 2 factorial of sex by F_0 by variation in F_2.

Analysis accuracy was evaluated by comparing the predicted results of the analysis with the observed results. Predicted results were obtained using linear regression to calculate frequency values for every 12.8-msec segment of each synthesized vowel. These values then served as comparative data for the computational procedures used for real speech.

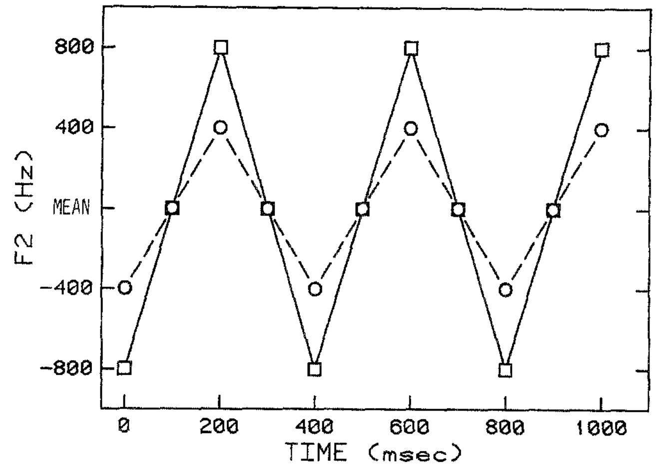

FIGURE 2. Variability of fundamental frequency for the synthetic vowel /ɑ/. The solid line represents 800-Hz variability and the dashed line represents 400-Hz variability around the mean F_2.

RESULTS

Validity of the Formant Frequency Analysis Method

Vowels having constant second formant frequencies. Table 1 shows the results of the formant frequency analysis of the four synthesized vowels having constant F_2 frequencies. The mean formant frequency fluctuation values were slightly greater than zero, showing a small amount of analysis error. The highest mean FFF value of 7.56 Hz, with a 5.65-FFFR value, was obtained for the synthesized vowel representing the female pathologic speaker with exaggerated variability of F_0, while the smallest mean FFF value of 1.55 Hz, with a 1.15-FFFR value, was obtained for the vowel representing the normal female speaker. Apparently, the greatest error resulted from the interaction of the acoustic characteristics of a female speaker and a varying F_0 rather than from either of these two factors individually.

TABLE 1. Mean frequency, frequency standard deviation, mean formant frequency fluctuation ($\bar{x}$FFF), and formant frequency fluctuation ratio (FFFR) derived from spectral analysis of four synthesized vowels.

Synthesized vowels		*Analysis results*			
Second formant frequency	*Fundamental frequency*	*Mean formant frequency*	*SD*	*$\bar{x}$FFF*	*FFFR*
1229	Constant	1227	2.13	1.67	1.36
1229	Varying	1226	4.42	3.84	3.13
1345	Constant	1338	2.07	1.55	1.15
1345	Varying	1337	23.07	7.56	5.65

Note. All values are in Hz.

TABLE 2. Predicted and observed results for the analysis of eight synthesized vowels having varying second formant frequencies.

Synthesized vowels			*Predicted results*				*Observed results*			
$\bar{x}F_2$	F_2 *variation around mean*	F_0	$\bar{x}F_2$	*SD*	$\bar{x}$*FFF*	*FFFR*	$\bar{x}F_2$	*SD*	$\bar{x}$*FFF*	*FFFR*
1229	400	Constant	1229	116.0	24.77	20.12	1232	116.3	25.72	20.87
1229	400	Varying	1229	116.0	24.77	20.12	1234	116.8	25.40	20.57
1229	800	Constant	1229	235.0	49.55	40.34	1230	232.4	50.03	40.64
1229	800	Varying	1229	235.0	49.55	40.35	1236	231.2	49.25	39.81
1345	400	Constant	1345	117.5	24.77	18.43	1368	125.2	27.92	20.40
1345	400	Varying	1345	117.5	24.77	18.43	1363	104.5	26.18	19.20
1345	800	Constant	1345	235.0	49.55	36.86	1369	221.4	51.55	37.65
1345	800	Varying	1345	235.0	49.55	36.86	1358	216.0	47.57	35.04

Note. All values are in Hz.

The median difference between the mean formant frequency derived from the analysis and the actual mean frequency of the synthesized vowel, another measure of analysis error, was 5 Hz, or a .38% discrepancy. The difference ranged from 2 Hz in the male speaker's vowel with a constant F_0 to 8 Hz in the female speaker's vowel with a varying F_0.

The degree of analysis error observed in this validity testing is far lower than that reported by Monsen and Engebretson (1983). The discrepant findings probably are related to differences in the vowels analyzed, their use of parallel rather than serial synthesis, and the fact that the formant frequency value for each vowel in the present study represented the average of 78 measurements, a large sample that would be less sensitive than a smaller sample to the effect of extreme values.

Vowels having varying second formant frequencies. The results for the eight synthesized vowels having varying F_2 frequencies are presented in Table 2, while the differences between observed and predicted values are presented in Table 3. The observed FFF values were all quite close to the predicted values. Table 3 shows the differences between observed and predicted values. The greatest FFFR difference was 1.97, obtained for the synthesized vowel representing the female pathological speaker with exaggerated variability of F_0 and F_2 frequency varying 400 Hz around the mean. The smallest FFFR difference, .30 Hz, was found for the vowel representing a normal male speaker with a constant F_0 and an 800-Hz variation in F_2. As in the analysis of the vowels having constant F_2 frequencies, the greatest error occurred in the analysis of vowels having the acoustic characteristics of a female speaker, although the parameter of constant or varying F_0 had a negligible effect. For all eight vowels, the median difference between observed and predicted values for $\bar{x}$FFF of 1.18 Hz, with a median difference in FFFR of .76, indicated a relatively small degree of error. It is interesting that the size of F_2 variation, that is, 400 Hz or 800 Hz, did not appear to affect the accuracy of analysis.

TABLE 3. Differences between observed and predicted results for the analysis of eight synthesized vowels having varying second formant frequencies.

Synthesized vowels			*Difference values*			
$\bar{x}F_2$	F_2 *variation around mean*	F_0	$\bar{x}F_2$	*SD*	$\bar{x}$*FFF*	*FFFR*
1229	400	Constant	3	.28	.95	.75
1229	400	Varying	5	.79	.63	.45
1229	800	Constant	1	2.63	.48	.30
1229	800	Varying	7	3.77	.30	.54
1345	400	Constant	23	7.76	3.15	1.97
1345	400	Varying	18	12.99	1.41	.77
1345	800	Constant	24	13.46	2.00	.79
1345	800	Varying	13	18.93	1.98	1.82

Note. All values are in Hz.

Results of the Formant Frequency Analysis

The analysis results for the five tardive dyskinetic subjects and the 10 normal subjects are presented in Table 4. For the pathologic subjects, the mean FFF ranged from 11.10 to 35.70 Hz, with a median of 33.04 Hz. Mean FFF for the normal subjects ranged from 2.86 to 13.60 Hz, with a median for all 10 subjects of 9.59 Hz.

The difference between the normal and pathological groups is striking; however, a comparison of results using only the mean FFF can be misleading. Since the absolute size of perturbation measures, such as the mean FFF, is dependent on the base quantity, comparisons of measures having different base quantities can lead to error. For example, Horii (1979) reported that since jitter size (the cycle-to-cycle perturbation of vocal periods) is highly correlated with fundamental frequency, many researchers express jitter in relation to the mean vocal period. The results of the present study show that mean FFF increases as the mean formant frequency increases (r = .95). Thus, the ratio of the mean formant frequency fluctuation to the mean formant frequency serves as a standardized representation of formant frequency fluctuation. This standardization method allows meaningful comparisons among vowels having different mean formant frequency values. Four of the five tardive dyskinetic subjects had ratio values that exceeded the mean of

TABLE 4. Characteristics of the second formant frequency from 1-sec segments of the vowel /ɑ/.

Subject	*Sex*	*Mean F_2 (Hz)*	*Frequency SD (Hz)*	*Mean formant frequency fluctuation (Hz)*	*Formant frequency fluctuation ratio*
Tardive dyskinetic					
1	F	1537	64.45	34.85	22.68*
2	F	1581	28.36	11.10	7.02
3	F	1731	140.16	33.04	19.08*
4	F	1104	46.14	35.70	32.33*
5	M	1253	92.76	19.40	15.48*
Normal					
1	M	1092	4.32	2.86	2.62
2	M	1061	9.90	7.27	6.85
3	F	1546	10.56	3.95	2.55
4	F	1259	16.60	13.60	10.80
5	M	1167	26.22	11.48	9.83
6	F	1342	11.67	10.53	7.85
7	M	1119	22.40	11.05	9.88
8	F	1318	10.10	5.60	4.25
9	F	1229	16.38	11.71	9.53
10	M	1064	14.82	8.64	8.13

*Formant frequency fluctuation ratio exceeds the mean of the normal subjects by two standard deviations.

the normal group ($\bar{x}$ = 7.23) by more than two standard deviations. The normal speakers' ratios ranged from 2.55 to 10.80, with a standard deviation of 3.08.

DISCUSSION

In this study formant frequency fluctuation analysis was able to measure the reduction in motor steadiness in four of five tardive dyskinetic patients who had been diagnosed as having abnormal involuntary movements of the orofacial structures. The FFFR values for these four tardive dyskinetic subjects were quite large, exceeding the mean of the normal group by two standard deviations. The other tardive dyskinetic subject had FFFR values that fell within the normal range.

The finding that one of the pathological subjects had normal results is not surprising in light of the results of a perceptual analysis of these subjects' speech production (Gerratt et al., Note 1). For the three articulatory dimensions of imprecise consonants, distorted vowels, and irregular articulatory breakdowns, this tardive dyskinetic subject was judged to be normal or only slightly abnormal. The FFF values for this subject were quite consistent with the perceptual judgments of articulatory adequacy. While having little articulatory disturbance, she was judged to be dysarthric primarily because of severe deviation from normal in the speech dimensions related to phonation. She was able to maintain normal steadiness of the articulatory structures during vowel production, but had little motor control of the laryngeal system.

The validity of the formant frequency fluctuation analysis was tested using synthesized vowels which had either constant or varying F_2 frequencies. Overall, the performance was reasonably good, with greater error observed for the vowels representing a female speaker. For the vowels having constant F_2 frequencies, one measure of error is the discrepancy between the actual mean formant frequency and that derived from the validity analysis. Across all four vowels, the median discrepancy of .38% shows that the method yields an acceptably small degree of inaccuracy in formant extraction. A more crucial measure of error is provided by the FFFR values. Since these vowels were synthesized with constant formant frequencies, the presence of formant frequency fluctuation of any size can be considered to be the result of analysis error. While the amount of error of analysis for the two vowels having nonvarying fundamental frequencies was rather small, the higher FFFR values for the two vowels having varying fundamental frequencies show a greater degree of error. In theory, the formant extraction technique estimates the resonance structure of the input spectrum while ignoring the fine spectral structure corresponding to the glottal source characteristics (Markel, 1972). However, these results show that in practice a varying F_0 does indeed have an effect on analysis stability.

In contrast to the results of the analysis of vowels having constant F_2 frequencies, the factor of F_0 variation did not appear to affect the accuracy of the analysis of vowels having varying F_2 frequencies. Looking at these eight vowels as four pairs in which each member of a pair differs only by F_0 (constant or varying), only one of the three pairs has a greater difference between observed and predicted FFFR values for the cognate having a varying F_0.

The two types of validity testing examined the accuracy of analysis in two different but complementary ways. One used constant F_2 frequencies to determine the stability of measurement over time, while the other evaluated the accuracy of measurement for F_2 frequencies that vary over time. Of the two varieties of validity testing, the accurate tracking of varying F_2 frequencies seems to be of greater importance since this task is the actual purpose of FFF measurement. Consequently, it was particularly assuring to find such a small overall degree of tracking error and no reduction of accuracy caused by a varying F_0, as found in the analysis of vowels having constant F_2 frequencies.

Nevertheless, the fact that the F_0 variation present in the vowels having constant F_2 frequencies apparently reduced the analysis stability must be considered in the interpretation of formant frequency fluctuation for dysarthric patients, since many patients having dysarthria often exhibit some form of vocal pathology (Darley et al., 1969a). Indeed, the dysarthria associated with tardive dyskinesia has been reported to include vocal harshness, strained-strangled voice, and voice tremor (Gerratt et al., Note 1).

The error for the vowel representing the male pathologic speaker having a varying F_0 was slightly larger than that for the vowel representing the male speaker with a nonvarying F_0, but was still acceptably small. On the other hand, the FFFR of 5.65 found for the vowel representing the female pathologic speaker is great enough for concern. In light of this degree of error, the results for the three female tardive dyskinetic subjects whose FFFR values were much larger than the mean of the normal group must be analyzed with some degree of caution. These subjects did indeed have a large amount of cycle-to-cycle variability in F_0, as reported by Gerratt, Fisher, and Brayton (Note 2). The median jitter ratio for the three female subjects was 93.87, while the median for the age-matched normal subjects was 8.37. Although it is impossible to determine the true amount of analysis error caused by an interaction of the glottal source characteristics and the formant extraction procedure, it must be assumed that some part of their FFFR values represents system error. However, the FFFR values for these subjects were so large that even when the amount of error is set at 5.65 for each—an amount equal to the error found in the validity test for the vowel representing a female pathologic speaker with constant F_2 frequencies—the remaining values are still greater than the mean of the normal group by two standard deviations. Nonetheless, it is clear that formant frequency fluctuation analysis using this type of LPC algorithm is somewhat less stable for female speakers with large degrees of variability in fundamental frequency.

Although other methods of formant extraction were not attempted, other techniques may be considered. Childers (1977) compared another linear prediction method, the closed-phase procedure, to the autocorrelation method used in the present study. A comparison of linear predictive spectra produced by these two methods was made on the vowel /ɑ/ spoken by a 60-year-old man whose epiglottis and two ventricular folds had been surgically removed and who had one paralyzed vocal cord. A difference of 20% in the estimated location of the second formant was found. Childers's interpretation of this discrepancy was that a source zero occurring within the vicinity of 1800 Hz in the spectrum produced by the autocorrelation method supressed the peak of the second formant, leaving shoulders on either side. One of the shoulders could have been picked mistakenly as F_2. This finding may indicate that the closed phase method is less sensitive to the glottal source characteristics than the autocorrelation method. Considering the vocal problems occurring in the various dysarthrias, the closed phase linear prediction analysis should be investigated for its potential use with this population.

Since all types of dysarthria include a reduction of motor control, this measure may also be useful as an index of the steadiness of the supraglottal vocal tract musculature in other varieties of motor speech disorders. For example, the effects of inaccuracy in the timing, range, force, and direction of articulatory movements often associated with ataxic dysarthria may also be analyzed objectively using this technique. While the determination of formant frequency fluctuation has the appealing advantage of providing one measure for the amount of all articulatory movement, it does not provide information about which of the articulators contributed to the overall fluctuation value. It is possible, for example, for an individual to attain the same formant frequency fluctuation value for different combinations of articulatory movement. However, the actual sequence of movements is of less clinical importance than the overall degree of involuntary articulatory movement.

ACKNOWLEDGMENTS

This research is based on a portion of a doctoral dissertation completed at Northwestern University. The author wishes to gratefully acknowledge the guidance provided by Hilda Fisher. Gratitude is expressed to Christopher Goetz for the neurological diagnosis of the subjects, and Gerald Canter and David Rutherford for their helpful advice. Appreciation is offered to Shaun Brayton, Douglas Noffsinger, and Charles Martinez for their valuable comments on the manuscript. Portions of this study were presented at the Annual Convention of the American Speech-Language-Hearing Association, Los Angeles, 1981. The research was supported in part by a National Research Service Award, Training Grant #5 T32 NS07100, from the National Institute of Neurological and Communicative Disorders and Stroke, National Institutes of Health.

REFERENCE NOTES

1. GERRATT, B. R., FISHER, H. B., & GOETZ, C. G. *A perceptual investigation of the dysarthria accompanying tardive dyskinesia.* Paper presented at the Annual Convention of the American Speech-Language-Hearing Association, Detroit, 1980.
2. GERRATT, B. R., FISHER, H. B., & BRAYTON, S. A. *Formant frequency fluctuation and jitter in the dysarthria associated with tardive dyskinesia.* Paper presented at the Annual Convention of the American Speech-Language-Hearing Association, Los Angeles, 1981.

REFERENCES

ATAL, B. S., & SCHROEDER, M. R. Predictive coding of speech signals (pp. 360-361). *Proceedings of the 1967 IEEE Conference on Communication and Processing,* 1967.

CHILDERS, D. G. Laryngeal pathology detection. *CRC Critical Reviews in Bioengineering,* 1977, *2,* 375-425.

DARLEY, F. L., ARONSON, A. E., & BROWN, J. R. Differential diagnostic patterns of dysarthria. *Journal of Speech and Hearing Research,* 1969, *12,* 246-269. (a)

DARLEY, F. L., ARONSON, A. E., & BROWN, J. R. Clusters of deviant speech dimensions in the dysarthrias. *Journal of Speech and Hearing Research,* 1969, *12,* 462-496. (b)

FANT, C. G. M. *Speech sounds and features.* Cambridge: MIT Press, 1973.

HORII, Y. Fundamental frequency perturbation observed in sustained phonation. *Journal of Speech and Hearing Research,* 1979, *22,* 5-19.

KLATT, D. H. Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America,* 1980, *67,* 971-995.

MARKEL, J. D. Digital inverse filtering—A new tool for formant trajectory estimation. *IEEE Transactions on Audio and Electroacoustics,* 1972, *20,* 129-137.

MARKEL, J. D., & GRAY, A. H. *Linear prediction of speech.* New York: Springer-Verlag, 1976.

MARSDEN, C. D., TARSY, D., & BALDESSARINI, R. J. Spontaneous and drug induced movement disorders in psychotic patients. In D. F. Benson & D. Blumer (Eds.), *Psychiatric aspects of neurologic diseases.* New York: Grune & Stratton, 1975.

MAXWELL, S., MASSENGILL, R., & NASHOLD, B. Tardive dyskinesia. *Journal of Speech and Hearing Disorders*, 1970, *35*, 33-36.

MONSEN, R. B., & ENGEBRETSON, A. M. The accuracy of formant frequency measurements: A comparison of spectrographic analysis and linear prediction. *Journal of Speech and Hearing Research*, 1983, *26*, 89-97.

PETERSON, G. E., & BARNEY, H. L. Control methods used in a study of the vowels. *Journal of the Acoustic Society of America*, 1952, *24*, 175-184.

PORTNOY, R. A. Hyperkinetic dysarthria as an early indicator of impending tardive dyskinesia. *Journal of Speech and Hearing Disorders*, 1979, *44*, 214-219.

SCHROEDER, M. R. Linear prediction, extremal entropy, and prior information in speech signal analysis and synthesis. *Speech Communication*, 1982, *1*, 9-20.

WAKITA, H. Instrumentation for the study of speech acoustics. In N. J. Lass (Ed.), *Contemporary issues in experimental phonetics*. New York: Academic Press, 1976.

Received October 20, 1981
Accepted July 29, 1982

Requests for reprints should be sent to Bruce Gerratt, Wadsworth VA Medical Center, Audiology and Speech Pathology Service (691/126), Wilshire & Sawtelle Boulevards, West Los Angeles, CA 90073.

 0022-4685/83/2602-0304$01.00/0