Original Paper

Folia Phoniatr Logop 2006;58:323–339
DOI: 10.1159/000094567

Folia Phoniatrica et Logopaedica

Fundamental Frequency as a Perceptual Cue for Vowel Identification in Speakers with Parkinson's Disease

Kate Bunton

Institute for Neurogenic Communication Disorders, University of Arizona,
Tucson, Ariz., USA

Key Words

Parkinson's disease · Vowel identification · Fundamental frequency · Speech perception

Abstract

This study investigates the importance of fundamental frequency (F0) as a perceptual cue for identification of vowel targets produced by speakers with Parkinson's disease (PD). It has been suggested in the literature that F0 is a redundant cue for vowel identification in highly intelligible speech. For speakers with dysarthria who are having difficulty with segmental and suprasegmental aspects of production which result in ambiguous or conflicting cues in the acoustic signal, F0 may have increased perceptual importance for accurate identification of vowel targets. In the present study, F0 contours for single-word targets produced in sentence level material by 20 speakers with PD and 20 control speakers were synthetically modified in several different ways (i.e., flattened and enhanced). Listener identification of vowel targets across the F0 conditions was recorded. The accuracy of vowel identification for the control group was not affected by the flattening of the F0 contour. For the speakers with PD, however, modification of the F0 contour (flattening or enhancing) affected the accuracy with which listeners identified certain vowels. Differences in vowel identification were found primarily for the front vowels /ɪ, ɛ, æ/ along a high-low continuum.

Introduction

Vowel formants and fundamental frequency (F0) vary over the course of a syllable, and it is not known which properties of this spectral-temporal variation are the primary perceptual cues for vowel identification. Most commonly, vowel identity is thought to be determined by the location of the first three formant frequencies

KARGER
Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

© 2006 S. Karger AG, Basel
1021–7762/06/0585–0323$23.50/0

Accessible online at:
www.karger.com/fpl

Kate Bunton, PhD, Institute for Neurogenic
Communication Disorders, Speech and Hearing Sciences
University of Arizona, PO Box 210071
Tucson, AZ 95721-0071 (USA), Tel. +1 520 621 2210
Fax +1 520 626 2226, E-Mail bunton@u.arizona.edu

(F1, F2, F3). F0 is believed to primarily influence perception of speaker-specific qualities (e.g., age, sex) [1–3]. These beliefs stem from several studies using synthetic speech where the effects of independent manipulations of formant and F0 transitions on vowel perception were examined [3–5]. In these studies, the effect of F0 shifts on vowel identification was modest relative to the effects of F1 and other formant frequencies.

Changes of greater than an octave in F0 were needed to shift vowel identification [3–5]. In contrast to the relatively small influence of F0 found in these studies, Traunmuller [6] reported significant F0 effects on the perception of some vowels. In particular, the F1–F0 Bark distance was reported to be a nearly invariant perceptual correlate of perceived vowel height F1–F0 for front vowels. This finding was confirmed by Syrdal [7] and Syrdal and Gopal [8] in analyses of two large data sets of American English vowels. Although recognizing the importance of F0 in vowel identification, several researchers have reported that for highly intelligible speakers F0 may be a redundant cue that is not necessary for vowel identification, at least in ideal conditions [9, 10]. In situations where listeners are presented with a less than ideal signal, however, they may be forced to alter their perceptual strategies and trade cues in the speech signal off against one another in order to reach a perceptual decision [11]. Use of a trading relation strategy is likely when one component of the speech chain is degraded (e.g., speaker or transmission environment). For example, when segmental cues for vowel identification are reduced or impaired, such as when the speaker is dysarthric, F0 may help the listener identify the intended segment. F0 has also been demonstrated to have increased importance in a poor transmission environment such as with background noise [12, 13].

In addition to segmental production deficits which affect the acoustic characteristics of segments, speakers with dysarthria, particularly of the type associated with Parkinson's disease (PD), may have deficits in suprasegmental production that decrease the range of F0 variability they produce. The interaction between segmental and suprasegmental production deficits and their combined effects on segment identification and overall speech intelligibility is not known.

The results of several studies have shown that artificially flattening the F0 contour of normal speakers, much like the F0 flattening that occurs in dysarthric speech, reduces their sentence intelligibility [12–15]. There are several possible explanations for these observations. First, there is evidence that listeners use suprasegmental information to assign an initial syntactic structure before decoding the rest of the message [16–19]. Another possible explanation for the reduction in speech intelligibility associated with flattened F0 contours may be that F0 has an effect on the identifiability of individual speech segments. Kent and Rosenbek [20] have reported that flattened F0 contours blur the contrast between adjacent units in the speech signal by reducing the salience of acoustic cues to segment identity.

Computer-implemented improvements in F0 contours for speakers with hearing impairments have been correlated with significant improvements in speech intelligibility scores [21–23]. Osberger and Levitt [23] report that the extent to which intelligibility was altered as a result of suprasegmental modifications was directly related to the frequency of segmental errors, particularly vowel segments. These researchers suggest that difficulties with production of prosodic variables coupled with the presence of segmental-level articulatory deficits may have provided listeners with conflicting perceptual cues and thus had a negative influence on a listener's ability to make perceptual decisions about the intended speech segments. Consistent with this claim, research examining speech production characteristics of speakers with dysarthria has reported that F2, which is a psychoacoustically and linguistically salient cue in normal speech production, may be less stable for speakers who may have segmental-level articulatory deficits [24]. Hence it may be less heavily weighted by listeners as a cue for segment identification. A more salient cue, such as F0, may become a relatively more important cue for vowel identification in such cases. This is a case of trading relations, where a listener alters his/her perceptual strategies to arrive at a given percept based on the information available (or lack of) in the acoustic signal [11].

The purpose of this study was to determine whether vowel identification is affected by changes in the F0 contour of speech produced by persons with dysarthria associated with PD, each of whom had difficulty with both segmental and suprasegmental aspects of production.

Method

Speakers

Acoustic data were collected from a total of 40 speakers, 20 diagnosed with PD and 20 with no history of neurologic disease who served as a control group (CG). Speakers were between the ages of 49 and 72 years, with an equal number of males and females in each group. For speakers with PD, a medical diagnosis was required, however, there were no explicit criteria regarding disease stage, duration, or manifestation. All speakers with PD were stabilized on anti-Parkinson medications and did not demonstrate drug-induced dyskinesia. Speakers were recorded 1–2 h after receiving their medication for PD. The primary criterion for inclusion in the study was that the speaker had an intelligibility deficit. Single-word estimates of speech intelligibility that fell in the range of 60–85% for speakers with PD were suitable for the current study. This range of intelligibility impairment ensured that speakers were likely to produce segmental-level vowel errors, a dependent variable in the study.

A short battery of screening measures was administered to all potential participants (PD and CG). To be eligible for participation, speakers needed to: obtain a score of at least 25/30 on the Mini-Mental Test [25] as a screen for dementia; report English as their first language; pass a hearing screening at 40 dB HL for frequencies of 0.5, 1, 2, and 4.0 kHz [26] and report no history of taking drugs known to produce oromotor dyskinesia (other than those currently being taken in association with PD). If a potential participant passed these screening items, a 70-item revision of the original intelligibility instrument by Kent et al. [27] was administered. Single words were chosen as a screening tool as they give a good indication of whether speakers were likely to produce segmental level errors during the experiment. The words were recorded on DAT tape (Tascam DA-P1) with a head mount microphone (Crown CM-311A) placed 3.5 cm in front of the speaker's mouth. The words were digitized and stored in CSpeech (filter cutoff 9.8 kHz, sampling rate = 22.1 kHz) [28]. The same software was used to present a randomized order of the words to listeners. To obtain estimates of speech intelligibility, single-word tokens from each speaker were presented over a loudspeaker at a comfortable listening level to 4 listeners who were seated individually in a sound-treated room. A response form with numbered rows was given to each listener and they were asked to provide a broad transcription of each word immediately after hearing it. A score of correctly heard words was generated for each listener. Single-word intelligibility scores for a given speaker were computed by averaging scores from the 4 listeners. No intelligibility criterion was used for speakers in the CG. For speakers with PD, if their score fell between 60 and 85% they were eligible for the study. A total of 31 speakers with PD were screened for the present study, 11 did not meet inclusion criteria based on their intelligibility scores. Individual speaker characteristics, including time since diagnosis and word intelligibility scores can be found in table 1.

Table 1. Individual speaker characteristics

Speaker code	Sex	Age	Diagnosis	Duration years	Intelligibility mean (SD)
PS01	F	62	PD	5	66.4 (2.7)
PS02	F	50	PD	4	73.9 (4.2)
PS03	F	68	PD	3	78.1 (3.5)
PS04	F	70	PD	3	83.4 (2.1)
PS05	F	72	PD	6	79.8 (8.4)
PS06	F	61	PD	6	83.6 (8.7)
PS07	F	66	PD	7	85.3 (4.6)
PS08	F	66	PD	9	80.6 (2.6)
PS09	F	60	PD	7	78.5 (2.5)
PS10	F	60	PD	6	80.6 (5.9)
PS11	M	72	PD	7	83.7 (5.23)
PS12	M	72	PD	3	71.5 (1.6)
PS13	M	66	PD	7	79.8 (2.5)
PS14	M	71	PD	8	67.8 (5.6)
PS15	M	72	PD	6	84.7 (3.4)
PS16	M	71	PD	9	71.6 (3.8)
PS17	M	70	PD	10	75.9 (6.1)
PS18	M	71	PD	10	78.3 (1.9)
PS19	M	64	PD	5	84.6 (5.1)
PS20	M	49	PD	1	83.5 (2.5)
NS01	F	70			99.4 (0.6)
NS02	F	68			98.7 (0.3)
NS03	F	68			96.5 (2.1)
NS04	F	66			100
NS05	F	67			100
NS06	F	49			97.4 (0.6)
NS07	F	62			97.6 (0.9)
NS08	F	59			100
NS09	F	60			100
NS10	F	70			97.8 (0.6)
NS11	M	53			100
NS12	M	57			97.6 (2.1)
NS13	M	64			98.5 (0.8)
NS14	M	54			96.5 (3.1)
NS15	M	68			98.4 (2.0)
NS16	M	63			96.8 (3.5)
NS17	M	67			97.4 (1.2)
NS18	M	57			100
NS19	M	66			98.5 (1.1)
NS20	M	63			97.8 (1.3)

The duration column indicates the time between diagnosis and recording. Blanks within the table indicate not applicable.

Listeners

Sixty listeners participated in the present study. Twenty listeners were used in the screening portion described above and the remaining 40 listeners were used in the main portion of the experiment. Listeners were recruited from the university community, and all had completed an introductory phonetics course and could perform broad transcription. This skill was necessary to complete the transcription task (see below). Additional inclusion criteria included: (1) between the ages of 18–50 years, (2) no experience with speakers with dysarthria and (3) able to pass a hearing screening at 25 dB HL for frequencies of 0.5, 1, 2, and 4.0 kHz [29].

Experimental Speech Sample

The speech sample consisted of 48 single-syllable words. There were 6 tokens for each of 8 English monophthong vowels, /i/, /ɪ/, /ɛ/ /æ/, /u/, /ʊ/, /ɔ/, and /ɑ/. Many of the tokens were taken from the Kent et al. [27] intelligibility test because they addressed one of three phonetic vowel contrasts: front vs. back, high vs. low, or long vs. short. A number of additional tokens were added so each target vowel was sampled equally. Because the listening task in the current experiment involved transcription, rather than multiple choice as in the Kent et al. [27] intelligibility test, the option of real-word foils for changes in vowel identification along these continua was not critical. The 48 target words are listed in 'Appendix A'. All single words were elicited in a meaningful sentence under normal speaking conditions. Sentences were between 6 and 8 syllables in length with the target word as a content word placed in the beginning or middle portion of the sentence. All target words received primary stress in the sentence. If a speaker did not produce primary stress on the target word, as judged by the investigator, a cue card with the target word italicized was presented. Speakers were then asked to repeat the sentence, placing stress on the italicized word. For example: 'We can *shoot* or do nothing'; 'That *bed* is too small'; 'There is a *leak* in this cup'.

Acoustic Measures

An automatic pitch tracker in CSpeech [28] was used to compute F0 contours for each sentence and speaker. Each F0 contour was visually inspected and manually corrected to eliminate spurious values (too high or low). The typical F0 range (mean, minimum, and maximum) for each sentence was recorded. The mean F0 was used during the flat F0 modification condition (see below). The minimum and maximum values were used to describe the typical use of F0 in sentence production for the two speaker groups. The minimum and maximum F0 values for the target word within each sentence were also recorded.

First and second formant frequencies from the four English corner vowels /i, æ, u, ɑ/ were measured at the temporal midpoint of the vowel targets. Measurements were made using both broadband spectrogram and linear predictive coding (LPC) displays. Mean formant values were used to create F1-F2 plots to define the vowel quadrilateral for each group. Geometric area values were calculated by summing the areas of the two triangles created by bisecting each quadrilateral [18, 19, 30, 31].

Construction of Stimuli

The vocalic segments within each sentence were marked in CSpeech by identifying the first and last glottal pulse using a spectrogram to guide placement of the cursors. An LPC synthesis algorithm in CSpeech [32] was then applied to the marked segments to meet one of the following conditions:

(1) *original F0,* a set of unmodified, resynthesized sentences was used as a control condition to ensure that any potential shifts in vowel identification were not due to the resynthesis process;

(2) *flat F0,* by setting the pitch period of all vocalic segments equal to the mean of the sentence;

(3) *enhanced F0,* an F0 contour from a single speaker considered to be representative of the contour used by the CG for that sentence was selected. This model F0 contour was superimposed on the sentences produced by the PD speakers. For each sentence, the vocalic segments were marked using cursors and the pattern of F0 movement was modified to match the model selected.

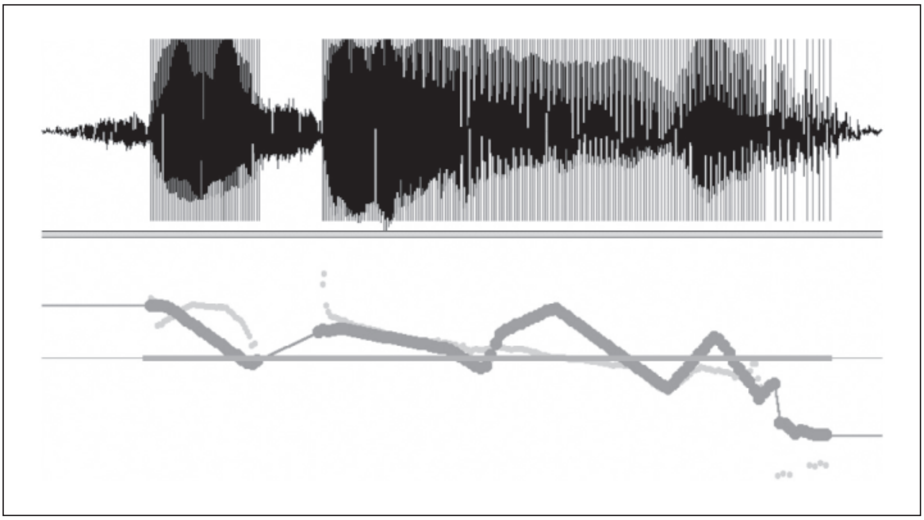

Fig. 1. Examples of glottal waveform and F0 contours for the utterance ‘She saw him do it’ produced by a female speaker with PD. The original resynthesized F0 contour is shown in light gray, the flat F0 contour is shown as a straight line, and the enhanced F0 contour as a black line.

For rising F0 segments, the initial value of the F0 contour was the same as in the original utterances, with the range extended by increasing the final value within the segment; for falling segments, the final value was the same as the original, with the range extended by increasing the initial value within the segment. In other words, the resynthesis procedure for each vocalic segment matched the shape of the model F0 contour (i.e., rises and falls), but was centered on each speaker’s habitual F0. A separate model contour was used for male and female speakers.

In summary, there were three F0 conditions for each speaker in the PD group and two F0 conditions for the CG. For the PD speakers, the F0 conditions included the original F0, flat F0, and enhanced F0 conditions. For the CG the conditions included the original F0 and flat F0 conditions. An example of the original, flat and enhanced F0 contours for the sentence ‘She saw *him* do it’ produced by a female speaker with PD can be seen in figure 1. In this figure the light gray line represents the original F0 contour produced by the speaker and the dark gray line represents the enhanced F0 contour. The flat, dark gray line represents the flattened F0 contour. F0 modifications were based on the entire sentence because of its potential to change the nature of the word-level F0 characteristics and because the sentence stimuli were used in complementary experiments. For this study, the target words were isolated, using spectrographic and auditory information to determine word boundaries, and saved to separate files to be used in the listening portion of the experiment.

Listening Task

A broad transcription task was selected to determine the accuracy with which the listener identified the intended speech segments within a target word. Forty listeners were seated individually in a sound-treated room and single words were presented over a loudspeaker placed 1 m in front of the listener. Words were presented in a background of 12-talker babble. A 30-second sample of babble was low-pass-filtered and digitized from the noise channel of a recording of the Speech Perception in Noise Test [33]. Digitization parameters were identical to those used for the speech stimuli. For each word presented, a segment of babble was started from a random location within this 30-second sample. The duration of this segment exceeded that of the target word so that it was centered temporally in the babble. Test words were presented at an overall level of

70 dB SPL with an S/B ratio of –10 dB. Ferguson and Kewley-Port [34] report that this S/B ratio prevents ceiling effects for the most intelligible vowels (i.e., highly intelligible speakers).

Each listener received written and oral directions for the task. They were asked to listen to each word and immediately perform a broad transcription of what they heard. The listeners were not given any information about the target words. They were instructed to be as accurate as possible in their transcription and were allowed to ask for a word to be repeated. Due to the large number of speakers and words in the experiment, listeners were randomly assigned to hear words from 5 speakers with PD and 5 from the CG. Each listener participated in two separate listening sessions with each session lasting no more than 30 min including a 5-min break. The words selected for presentation were randomized across speakers (CG and PD) and F0 condition (original F0, flat F0, and enhanced F0). An 8-second silent interval was used between words to allow the listeners time to respond. At the conclusion of the listening portion of the experiment, transcription from 10 listeners had been completed for each of the single-word targets produced by each speaker in each of the F0 conditions.

Statistical Analysis

The Fisher exact test of significance was used to detect differences in vowel identification across F0 conditions and vowel targets for individual speakers. This test was selected because of the low cell counts expected (i.e., a maximum of 8 based on the number of productions of each target vowel) [35]. Cell counts were based on listener identification of the vowels in each of the F0 conditions. Counts of the number of tokens identified as each vowel were collapsed into either 3 × 2 or 3 × 3 tables depending on the number of vowels identified. A significance value of 0.05 was used for all tests.

Reliability

Formant and fundamental frequencies were remeasured for 30% of the vowels produced by speakers with PD by the investigator and a second investigator to obtain estimates of inter- and intrajudge reliability. The overall correlation coefficient for intrajudge reliability was 0.91 across measures and speakers, for estimates of interjudge reliability the value was 0.89. Reliability was judged to be acceptable for measures of F0, F1, and F2. Ninety-five percent of the differences between original and second measures of F0 were less than 15 Hz, 85% of the differences between the original and second measures of F1 less than 30 Hz and 89% of the differences between original and second measures of F2 were less than 45 Hz. Inter- and intrajudge differences which exceeded these amounts were reassessed and modified as appropriate.

Measures of inter- and intralistener reliability were determined using Kendall's coefficient of concordance. The value of the coefficient for interlistener reliability was $W = 0.954$, indicating a high agreement between listeners. To obtain a measure of intralistener reliability, 150 tokens from the original set of 1,440 presented to each listener were heard twice. The value of the coefficient was $W = 0.872$. Kruskal-Wallis ANOVA by ranks showed no significant order effect on overall intelligibility for listeners ($H = 1.36$, $p = 0.879$), who had heard a randomly selected set of 5 speakers with PD and 5 from the CG.

Results

Acoustic Measures

Mean F0 values and standard deviations across the 48 sentences produced by each speaker can be seen in table 2. The mean F0 for individual speakers was age- and sex-appropriate. The mean for female speakers in the PD group was 210.8 Hz. The CG female speakers had a mean F0 of 224.2 Hz. For the male speakers, mean F0 values were 122.42 Hz and 119.8 Hz for the PD and CG, respectively. Mean sentence F0 variability for each speaker is also shown in table 2. This was calculated by subtracting the minimum F0 value from the maximum F0 value for each sentence

Table 2. Mean F0 and standard deviation (shown in parentheses), F0 variability for the sentence and target words and vowel space areas for individual speakers

Speaker code	Mean F0	F0 variability (sentence)	F0 variability (word)	Vowel space area, kHz^2
PS01	234.1 (4.2)	57.1 (13.1)	6.5 (2.0)	295
PS02	211.0 (6.8)	86.9 (9.7)	7.1 (3.5)	312
PS03	225.2 (13.5)	58.3 (16.5)	6.7 (1.6)	295
PS04	212.3 (21.2)	46.5 (18.7)	5.4 (1.5)	307
PS05	226.7 (7.9)	52.5 (16.4)	4.6 (2.1)	288
PS06	223.2 (11.4)	54.5 (5.4)	2.5 (1.7)	291
PS07	189.2 (9.8)	72.9 (11.9)	6.9 (4.1)	270
PS08	204.5 (22.4)	40.7 (8.7)	8.2 (3.1)	281
PS09	208.2 (5.6)	44.9 (6.7)	1.7 (0.9)	269
PS10	192.0 (11.4)	46.6 (16.4)	1.9 (1.0)	296
PS11	138.5 (14.8)	81.6 (14.5)	5.4 (2.1)	242
PS12	117.3 (16.5)	48.3 (13.3)	4.9 (1.2)	238
PS13	95.2 (14.6)	21.5 (6.1)	2.4 (1.1)	191
PS14	134.2 (15.9)	32.1 (13.5)	0.9 (0.1)	269
PS15	133.3 (12.3)	34.2 (8.9)	1.1 (0.5)	245
PS16	123.5 (11.9)	54.1 (14.3)	4.7 (2.8)	274
PS17	125.8 (15.6)	37.2 (15.4)	3.9 (2.6)	261
PS18	115.5 (9.8)	39.6 (9.8)	0.8 (0.2)	289
PS19	122.9 (16.2)	71.3 (14.2)	6.1 (4.7)	238
PS20	118.0 (13.4)	37.8 (8.1)	1.0 (0.6)	240
CG01	192.5 (11.1)	113.6 (11.2)	21.4 (6.8)	465
CG02	250.1 (15.4)	123.4 (14.5)	16.7 (9.7)	442
CG02	203.5 (13.5)	132.2 (16.3)	18.0 (4.0)	416
CG03	189.2 (13.1)	94.8 (15.8)	19.1 (10.1)	486
CG04	223.2 (9.8)	104.7 (14.6)	16.4 (9.2)	504
CG06	256.2 (16.1)	87.8 (21.5)	14.8 (6.4)	466
CG07	246.1 (15.2)	73.5 (22.4)	11.4 (3.1)	498
CG08	208.5 (16.2)	104.5 (17.1)	13.3 (5.4)	512
CG09	224.6 (14.2)	94.9 (13.1)	16.1 (2.9)	480
CG10	217.2 (16.5)	97.9 (17.2)	16.9 (8.7)	496
CG11	98.7 (14.6)	101.6 (11.7)	19.4 (8.2)	319
CG12	107.6 (8.2)	86.5 (15.5)	23.7 (11.1)	365
CG13	86.8 (9.1)	92.5 (118.9)	14.9 (8.9)	372
CG14	128.3 (10.4)	51.5 (114.5)	21.1 (10.5)	313
CG15	146.8 (11.2)	168.5 (16.9)	17.4 (13.0)	328
CG16	124.8 (11.0)	112.6 (8.3)	17.1 (5.6)	342
CG17	135.9 (7.9)	106.5 (17.8)	14.1 (4.7)	304
CG18	141.2 (16.5)	76.4 (14.6)	11.0 (6.9)	298
CG19	111.9 (14.6)	96.0 (6.9)	13.1 (3.8)	301
CG20	117.8 (8.4)	113.2 (11.7)	12.8 (7.5)	316

and then averaging across all 48 sentences produced by each speaker. The range of F0 variation for the CG (73–132 Hz for females and 40–113 Hz for males) was consistent with typical F0 ranges reported for simple declarative utterances (70–150 Hz) [36, 37]. For speakers in the PD group, the range was about half that of the controls

Table 3. Confusion matrix for vowel identification in the original F0 condition

	/i/	/ɪ/	/ɛ/	/æ/	/u/	/ʊ/	/ɔ/	/ɑ/
/i/	240							
/ɪ/		50	16					
/ɛ/		166	168	75				
/æ/		24	56	165				
/u/					240			
/ʊ/						227		
/ɔ/						13	240	
/ɑ/								240

The column labels indicated the intended target vowel produced by the speaker and the row labels indicate the vowel that was identified by the listener. Blank cells within the table indicate zero responses in that category.

Table 4. Confusion matrix for vowel identification in the flat F0 condition

	/i/	/ɪ/	/ɛ/	/æ/	/u/	/ʊ/	/ɔ/	/ɑ/
/i/	240							
/ɪ/		14	9					
/ɛ/		150	126	47				
/æ/		76	106	193				
/u/					240			
/ʊ/						227		
/ɔ/						13	240	
/ɑ/								240

The column labels indicated the intended target vowel produced by the speaker and the row labels indicate the vowel that was identified by the listener. Blank cells within the table indicate zero responses in that category.

(21–81 Hz for females and 51–68 Hz for males). Group differences in sentence level F0 variability were found to be statistically significant based on a two-sample t test ($t = 8.742$, $p = 0.000$). Group difference in the mean F0 variability for the target word was also statistically significant ($t = 7.453$, $p = 0.000$). The mean F0 variability (calculated the same as sentence F0 variability) for the target words ranged from 11.0 to 23.7 Hz for the CG and 0.9 to 6.9 Hz for the PD group (table 2).

Vowel quadrilaterals derived from the English corner vowels /i, æ, u, ɑ/ for the speakers with PD were smaller than those of the control group (table 2). For the CG, the mean vowel space area for females was 476.5 kHz^2 and for males was 325.8 kHz^2. The range for the group was 301–512 kHz^2. The mean vowel space area for the PD group was 248.7 kHz^2 for males and 290.4 kHz^2 for females. Areas ranged from 191 to 312 kHz^2. With the exception of 2 speakers, the calculated vowel area for the PD group was smaller than that of the CG.

Transcription Data

To determine listener identification of a vowel target, 8 of 10 listeners had to agree on its identification. This level of agreement was successfully applied to all words and speakers. For speakers in the control group, 98.4% of the target vowels were identified as the intended target in both F0 conditions (original and flat). For 1 female speaker (CG07), several words with the target vowel /ɪ/ were identified as /ɛ/ in the flat F0 condition. Target words were presented with a background of multi-speaker babble to prevent ceiling effects. The inclusion of noise, however, did not prevent ceiling effects for vowel identification in the present experiment. This finding is considered further in the discussion. For speakers in the PD group, changes in vowel identification were observed across the three F0 conditions. Confusion matrices showing listener identification of the target vowels can be seen in tables 3, 4, and 5 for the three F0 conditions (original, flat, and enhanced, respectively). The total in each column equals 240 (6 tokens per vowel × 20 speakers). The target vowels are listed across the columns and listener identification of the vowels is listed in rows. Results for each F0 condition are described below.

Original F0 Contour

In the original F0 condition, 36.5% (350 of 960) of the words produced by speakers with PD were identified as vowels other than the intended target vowel (hereafter referred to as errors) (table 3). The majority of errors (337 of 350) were recorded for the front vowels along a high-low continuum. The most frequent error was when a higher target vowel was identified as lower vowel. For example, of the 240 words where the target vowel was /ɪ/ only 50 were identified correctly. One hundred and sixty-six vowels were identified as the lower vowel /ɛ/ and 24 were identified as /æ/. Overall, 79% of the words with the target vowel /ɪ/ were identified by listeners as containing a lower front vowel. For words containing the target vowel /ɛ/, 56 were identified as the lower vowel /æ/ and 16 were identified as the higher vowel /ɪ/. In the case of the target vowel /æ/, 75 tokens were identified as the higher vowel /ɛ/. For the back vowels, there were few errors. Thirteen words where the target vowel was /ʊ/ were identified by listeners as /ɔ/. These errors were associated with 2 male speakers in the PD group. The target vowels /i/, /u/, /ɔ/, and /ɑ/ were never misidentified in the original F0 condition.

Flat F0 Contour

The overall error rate for the flat F0 condition was 41.7% (401 total errors) (table 4). Of these errors, there were 173 tokens where identification of the target vowel in the flat F0 condition was different than in the original F0 condition. This included 86 words where the vowel targets were identified accurately in the original F0 condition but were counted as errors in the flat F0 condition. The pattern of errors was similar to that seen in the original F0 condition, with the most frequent error being a higher vowel target identified as a lower vowel. For words with the target vowel /ɪ/, the number of errors was 226. Of these errors, 151 vowels were identified as /ɛ/ and 76 were identified as /æ/. For the target vowel /ɛ/, 106 of the vowels were identified as the lower vowel /æ/. This is nearly double the number recorded in the original F0 condition (n = 56). For the target vowel /æ/, more vowels were correctly identified in the flat F0 condition than in the original F0 condition (193 vs. 165). There were no errors recorded for the target vowel /i/ or the back vowels in the flat F0 condition.

Table 5. Confusion matrix for vowel identification in the enhanced F0 condition

	/i/	/ɪ/	/ɛ/	/æ/	/u/	/ʊ/	/ɔ/	/ɑ/
/i/	240							
/ɪ/		155	92					
/ɛ/		84	144	84				
/æ/		1	4	156				
/u/					240			
/ʊ/						240	1	
/ɔ/							239	
/ɑ/								240

The column labels indicated the intended target vowel produced by the speaker and the row labels indicate the vowel that was identified by the listener. Blank cells within the table indicate zero responses in that category.

Enhanced F0

The total number of errors counted in the enhanced F0 condition was 266 (table 5). This corresponds to an overall error rate of 36.5%. Vowel identification was different for 256 tokens compared to the original F0 condition. One hundred and eighty-one of these tokens had been 15 counted as errors in the original F0 condition but were accurately identified as the target vowel in the enhanced F0 condition. The pattern of vowel errors observed in this condition was opposite that seen in the original and flat F0 conditions. Targets which had been identified as low vowels in the original F0 condition were identified more frequently as higher vowels in the enhanced F0 condition. Often, the identification in the enhanced F0 condition matched the intended target vowel. For example, the target vowel /ɪ/ was correctly identified only 50 of 240 times in the original F0 condition, but was correctly identified 155 of 240 times in the enhanced F0 condition. The target vowel /ɛ/, on the other hand, was more frequently identified as /ɪ/ in the enhanced F0 condition (an error) compared to the original F0 condition (92 vs. 16). Only 4 tokens where the intended vowel was /ɛ/ were identified as the lower vowel /æ/ in the enhanced F0 condition compared to 56 times in the original F0 condition. The low vowel /æ/ which had been correctly identified with greater than 70% accuracy in both the original and flat F0 conditions was identified as the higher vowel /ɛ/ in 84 tokens in the enhanced F0 condition. No changes in vowel identification accuracy were seen for /i/ or the back vowels.

Identification Rates across F0 Conditions

The Fisher exact test was used to determine if the patterns of vowel identification for each speaker group were significantly different across the F0 conditions. Over 98% of the target vowels were identified as the intended vowel in both F0 conditions for the CG, so no statistically significant differences were found. For speakers in the PD group, the Fisher exact test was completed for each speaker and vowel target. No statistically significant differences were found between the three F0 conditions for the vowels /i/, /u/, /ʊ/, /ɔ/, and /ɑ/. For vowels where differences in listener identification were recorded for the three F0 conditions (i.e., /ɪ/, /ɛ/, and /æ/), results were statistically significant ($p < 0.05$) for many of the speakers. Results of the Fisher exact test for the target vowel /ɪ/ were statistically significant for 18 of 20 speakers. Differences were statistically significant for all 20 speakers for the target vowel /ɛ/, and for 13 of 20 speakers for the target vowel /æ/.

Discussion

The present experiment was designed to determine whether F0 variation had an effect on vowel identification in single words for persons with dysarthria associated with PD. Data were compared to a group of age-matched control speakers. For the CG, flattening the F0 contour did not have an effect on the accuracy of vowel identification. For speakers with PD, however, modification of the F0 contour (flattening or enhancing) did have a significant impact on vowel identification. This was particularly true for the front vowels /ɪ, ɛ, æ/ along a high-low continuum. These findings are consistent with previous research suggesting that F0 may be an important perceptual cue for vowel identification when the speech signal is degraded [12, 13, 15].

Effects of Background Noise

Speakers in the CG were judged to be highly intelligible based on their single-word intelligibility scores (table 1). To control ceiling effects in the present experiment, target words were presented in a background of multi-talker babble. Multi-talker babble was selected because several studies have reported that it has a greater negative effect on the identification of speech compared to that heard with a flat spectrum noise competitor [34, 38–40]. In the present study, however, the inclusion of background noise had very little effect on the accuracy of vowel identification. Vowels were identified with greater than 98.4% accuracy in both F0 conditions. The lack of interference from the background noise contradicts previous reports [15, 17, 34]. There are several differences between the present studies and those previously reported that are noteworthy. First, the speech stimuli used in the present study were single words. It has been suggested that listeners may have difficulty picking out the relevant perceptual information for segment identification in a background of multi-talker babble because of linguistic interference [41]. Because only single words were targeted in the present study, it was not necessary for listeners to track a linguistic message. Therefore, interference from background noise may have been minimal. Second, the present study reports only vowel identification accuracy whereas previous studies have reported overall speech intelligibility scores [15, 17, 34]. Based on (unreported) preliminary analysis of speech intelligibility for the data in the present study, it has been observed that listeners made many errors in consonant identification. Thus, it appears that ceiling effects were indeed controlled with a background of multi-talker babble.

To be consistent with the method of stimulus presentation used for the CG, speech stimuli for the PD group were also presented with background noise. It is possible, however, that the multi-talker babble could have interacted differently with this degraded speech signal and influenced listeners' identification of the vowels. To answer this question, a second (unreported) experiment where the speech stimuli from the PD group were presented to a different group of listeners in a quiet listening environment was completed. Differences in vowel identification between the two listening conditions (background noise vs. quiet) were minimal. Only 37 of 1,440 vowels (2.6%) were identified differently in the two conditions. All 37 were confusions between /ɪ/ and /ɛ/; 7 occurred in the flat F0 condition and 30 in the enhanced F0 condition. Therefore, it does not appear that inclusion of the background noise affected vowel identification accuracy rates for speakers in the PD group.

Vowel Identification

Statistically significant differences in listener vowel identification were observed across the three F0 conditions for the majority of speakers with PD. In the original F0 condition, 36.5% of vowels transcribed were counted as errors (i.e., identified as something other than the intended target vowel). This rate increased to 41.7% in the flat F0 condition, and decreased to 27.7% in the enhanced F0 condition. Differences in vowel identification were noted primarily for the three front vowels, /ɪ/, /ɛ/, and /æ/. Changes in identification were fairly systematic along a high-low continuum as the F0 contour was manipulated. In the original F0 condition, the direction of the errors was that a higher vowel target was identified as a lower vowel (e.g. /ɛ/ for /ɪ/ or /æ/ for /ɛ/). An identical pattern was seen in the flat F0 condition. In both of these conditions there was limited (or no) variability in the F0 contour for the vowel nucleus compared to that produced by the CG (table 2). In the original F0 condition this was the result of difficulty with suprasegmental production related to PD. In the flat F0 condition, the contour was purposely set to a mean F0 value for each speaker and essentially simulated a severe impairment in suprasegmental production. Even though speakers had reduced F0 ranges during habitual production, further flattening of the F0 contour resulted in an increase in the frequency of errors in segmental identification. There were 86 tokens which had been identified accurately in the original F0 condition that were counted as errors in the flat F0 condition (tables 3, 4). The high error rates in both F0 conditions can at least partially be explained by the F1–F0 difference that has been reported to be a cue for perception of vowel height [6–8]. Listeners likely integrate the time varying relation between F1 and F0 over the course of a segment to determine vowel identity [3]. A flat (or nearly flat) F0 contour, however, would have a fixed distance between F1 and F0 for the entire vocalic segment. Because there is no variation in this distance it may be perceived as larger than when the F0 contour is varied over the course of the segment. Vowels would, therefore, likely be perceived as being lower vowels than was actually intended. Enhancing the F0 contour, on the other hand, created a time varying F0 contour which effectively decreased the perceptual distance between F1 and F0. This should correspond to a shift in perception from a low vowel to a higher vowel [6–8] and is consistent with what was observed in the present study. Interestingly, in the enhanced F0 condition, many of the target vowels which had been counted as errors in the original F0 condition were identified accurately as the target vowel. For example, in the original F0 condition there were 190 errors recorded for the target vowel /ɪ/ (table 3), whereas in the enhanced F0 condition the number of errors decreased to 105 (table 5). A similar decrease in error rate was found for the target vowel /ɛ/ (table 5). Segmental-level production difficulties for speakers with PD potentially created ambiguous or conflicting cues in the acoustic signal which led to misidentification of the intended vowel in the original F0 condition. In other words, these vowels may have fallen near the boundary of the intended vowel category based on their acoustic characteristics, and therefore, modification of one of the cues

(F0) used by listeners to identify the segment had a significant effect on vowel identification. Evidence of difficulty with vowel articulation was observed in the reduced acoustic vowel space area for speakers in the PD group compared to the CG. Vowel space areas were approximately 50% smaller for the speakers with PD compared to the CG. Similar decreases in vowel space area for speakers with PD have been reported previously [18, 19, 42]. Further acoustic analysis of individual words in the present study would provide a more complete description of vowel characteristics, and whether or not there were indeed ambiguous or conflicting acoustic cues in the signal.

In addition to difficulty with segmental production, speakers with PD had simultaneous difficulty with suprasegmental aspects of production (table 2). Research on synthetic speech suggests that as the segmental structure of an utterance is degraded, the suprasegmental structure becomes more important for intelligibility [43, 44]. Consequently, target vowels which did not contain the expected acoustic signal characteristics and whose boundaries may be blurred due to the lack of F0 variation may have been ‘poor exemplars’ of the intended vowel category. In such cases, listeners were forced to adjust their perceptual strategies by readjusting the weights of the cues that contribute to identification of a segment. There were 78 tokens where vowel identification was different in all three F0 conditions. These tokens were likely not good exemplars of any particular vowel category and manipulation of one of the more salient perceptual cues used by listeners caused a shift in vowel identification. The shift in vowel identification for all of these tokens was in the expected direction given the F0 modification. Findings from the present study suggest that it is the combined effects of segmental and suprasegmental impairments that lead to errors in vowel identification as listeners are forced to adjust the weight of cues in the acoustic signal to optimize their perceptual strategies. Similar findings have been observed in the speech of people who are deaf [21–23, 45]. For highly intelligible speakers in the CG, on the other hand, acoustic characteristics of each vowel category were distinct and F0 was likely a redundant cue, therefore, manipulation of F0 did not affect the accuracy of vowel identification.

Variable Effects of F0 Manipulation on Vowel Identity

Differences in the effects of modification of the F0 contour were seen for the eight vowels sampled. Manipulation of the F0 contour had a significant effect on vowel identification for the three front vowels /ɪ, ɛ, æ/. No effects of F0 manipulation were recorded for the remaining vowels. Similar differences in vowel identifiability as a result of F0 manipulation have been reported previously [4, 5, 46–48]. In these studies, three vowels, /i/, /u/, and /ɑ/, had very low error rates when F0 was manipulated. These three vowels are English corner vowels, and thus represent extremes in the vowel space. Listeners are likely able to adjust to alterations in the vowel (e.g., manipulation of F0) without it affecting the vowel’s identity because of the contrastivity of other acoustic features. The corner vowel /æ/ has also been reported to have a low error rate as F0 was manipulated [4, 5, 46–48]. In the present study, however, errors for the vowel /æ/ were recorded in both the original and flat F0 condition. This is most likely related to underlying difficulties with production. It is not clear why the other corner vowels were not similarly affected since vowel space area was compressed for all speakers with PD compared to the CG (table 2). A nonsymmetric compression of vowel space has been reported previously for speakers with PD [42, 49]. This study reported a particular reduction in F2 for the vowel /æ/. The collapse of the vowel space combined with manipulations of F0 used in the present study likely led to the errors in identification of this vowel. Shifts in listeners' identification of the mid-front vowels /ɪ/ and /ɛ/ following manipulation of the F0 contour have been reported previously in the literature [46–48]. For these two front vowels, there are minimal differences along the F3-F2 dimension, which represents vowel place of articulation [8]. Thus the primary distinction between these vowels lies along the high-low or F1-F0 dimension. In contrast, the mid-back vowels /ʊ, ɔ/ are easily distinguished along both the F3-F2 and F1-F0 dimensions, therefore may have been more resistant to shifts in identification as F0 was manipulated.

Clinical Implications

The ultimate goal of research on speech intelligibility is to determine which specific changes can be made to increase intelligibility for a speaker. Based on results of the current study, it appears that F0 is an important perceptual cue for identification of vowel segments when speakers are experiencing segmental-level articulatory difficulties. The increased accuracy of vowel identification in the enhanced F0 condition compared to the original F0 condition is of particular interest. There are several treatment protocols for speakers with PD which focus on improving F0 variability [50–55]. If these protocols are successful in improving F0 variation, it is likely that the vowel identification accuracy will improve and this should result in improved speech intelligibility [24]. This hypothesis has not been tested directly. Application of this protocol to other speaker populations (e.g., ALS) is also of interest because of potential differences in segmental and suprasegmental production deficits related to the pathophysiology of the underlying disease and their potential combined effects on segment identification.

Acknowledgments

This work was supported by NIH R03 DC005902 from the National Institute on Deafness and Other Communication Disorders. Parts of this work were included in a presentation made at the 12th Biennial Conference on Motor Speech, 2004. I would also like to thank Gary Weismer, Kris Tjaden, and an anonymous reviewer for comments on an earlier version of the manuscript.

Appendix

1. beat
2. feed
3. feet
4. geese
5. leak
6. seed
7. bill
8. fill
9. him
10. lip
11. pit
12. sip
13. blend
14. fed
15. head
16. red
17. said
18. sell
19. bad
20. cash
21. fat
22. had
23. hat
24. pat
25. hoot
26. fruit
27. knew
28. moon
29. shoot
30. two
31. book
32. foot
33. full
34. hood
35. nook
36. put
37. cloth
38. dog
39. hall
40. haunt
41. wash
42. yawn
43. chop
44. dock
45. knot
46. lock
47. rock
48. top

References

1 Fujisaki H, Kawashima T: The roles of pitch and higher formants in the perception of vowels. IEEE Trans Audiol Electroacoust 1968;A11–16:73–77.

2 Ainsworth WA: Intrinsic and extrinsic factors in vowel judgments; in Fant G, Tatham M (eds): Auditory Analysis and Perception of Speech. London, Academic Press, 1975, pp 103–113.

3 Nearey T: Static, dynamic, and relational properties in vowel perception. J Acoust Soc Am 1989;85: 2088–2113.

4 Ryalls JH, Lieberman P: Fundamental frequency and vowel perception. J Acoust Soc Am 1982;72: 1631–1634.

5 Hirahara T, Kato H: The effect of F0 on vowel identification; in Tohkura Y, Vatikiotis-Bateson E, Sagisaka Y (eds): Speech Perception, Production and Linguistic Structure. Tokyo, Oshmusha, 1992, pp 89–112.

6 Traunmuller H: Perceptual dimensions of openness in vowels. J Acoust Soc Am 1981;69:1465–1475.

7 Syrdal A: Aspects of a model of auditory representation of American English vowels. Speech Commun 1985;4:121–135.

8 Syrdal A, Gopal H: A perceptual model of vowel recognition based on the auditory representation of American English vowels. J Acoust Soc Am 1986;79:1086–1100.

9 Cutler A, Dahan D, van Donselaar W: Prosody in the comprehension of spoken language: a literature review. Lang Speech 1997;40:141–201.

10 Stevens K: Diverse acoustic cues at consonantal landmarks. Phonetica 2000;57:139–151.

11 Repp BH: Phonetic trading relations and context effects: new experimental evidence for a speech mode of perception. Psychol Bull 1982;92:81–110.

12 Laures J, Bunton K: The effect of flattened fundamental frequency contours on sentence intelligibility. J Commun Disord 2003;36:449–464.

13 Bender B: Contributions of Fundamental Frequency to Speech Intelligibility; unpublished doct diss University of Memphis, Memphis, 2001.

14 Laures J, Weismer G: The effects of a flattened fundamental frequency on intelligibility at the sentence level. J Speech Lang Hear Res 1999;42:1148–1156.

15 Wingfield A, Lombardi L, Sokol S: Prosodic features and the intelligibility of accelerated speech: syntactic versus periodic segmentation. J Speech Lang Hear Res 1984;27:128–134.

16 Cutler A, Norris D: The role of strong syllables in segmentation for lexical access. J Exp Psychol 1988; 14:113–121.

17 Price P, Levitt A: The relative roles of syntax and prosody in perception of the /s/-/c/ distinction. Lang Speech 1983;26:291–304.

18 Liss J, Spitzer S, Caviness J, Adler C, Edwards B: Lexical boundary error analysis in hypokinetic and ataxic dysarthria. J Acoust Soc Am 2000;107:3415–3424.

19 Liss J, Spitzer S, Caviness J, Adler C, Edwards B: Syllabic strength and lexical boundary decisions in the perception of hypokinetic dysarthric speech. J Acoust Soc Am 1998;104:2457–2466.

20 Kent R, Rosenbek J: Prosodic disturbance and neurologic lesion. Brain Lang 1982;15:259–291.

21 Maassen B, Povel D: The effect of correcting fundamental frequency on the intelligibility of deaf speech and its interaction with temporal aspects. J Acoust Soc Am 1984;76:1673–1681.

22 Maassen B, Povel D: The effect of segmental and suprasegmental corrections on the intelligibility of deaf speech. J Acoust Soc Am 1985;78:877–886.

23 Osberger MJ, Levitt H: The effect of timing errors on the intelligibility of deaf children's speech. J Acoust Soc Am 1979;660:1316–1324.

24 Weismer G, Martin R: Acoustic and perceptual approaches to the study of intelligibility; in Kent R (ed): Intelligibility in Speech Disorders: Theory Measurement and Management. Amsterdam, Benjamin, 1992, pp 67–118.

25 Folstein M, Folstein S, McHugh P: 'Mini-mental state': a practical method for grading the cognitive status of patients for the clinician. J Psychiatr Res 1975;12:189–198.

26 Morrell C, Gordon-Salant S, Pearson JD, Brant LJ, Fozard JL: Age-and gender-specific reference ranges for hearing level and longitudinal changes in hearing level. J Acoust Soc Am 1996;100:1949–1967.

27 Kent R, Weismer G, Kent J, Rosenbek J: Toward phonetic intelligibility testing in dysarthria. J Speech Hear Disord 1989;54:482–489.

28 Milenkovic P: CSpeech for Windows (computer program). University of Wisconsin-Madison, Department of Electrical and Computer Engineering, 2000.

29 American Speech-Language-Hearing Association Panel of Audiologic Assessment: Guidelines for Audiologic Screening. Rockville, 1997.

30 Fourakis M: Tempo, stress, and vowel reduction in American English. J Acoust Soc Am 1991;90: 1816–1827.

31 Turner G, Tjaden K, Weismer G: The influence of speaking rate on vowel space and intelligibility for individuals with amyotrophic lateral sclerosis. J Speech Hear Res 1995;38:1001–1013.

32 Milenkovic P: F0Ramp (computer program). University of Wisconsin-Madison, Department of Electrical and Computer Engineering, 2002.
33 Kalikow D, Stevens K, Elliot L: Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. J Acoust Soc Am 1977;47:613–617.
34 Ferguson S, Kewley-Port D: Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. J Acoust Soc Am 2002;112:259–271.
35 Siegel S: Nonparametric Statistics for the Behavioral Sciences. New York, McGraw-Hill, 1956.
36 Kim HH: Monotony of speech production in Parkinson's disease: acoustic entities and their perceptual relations; unpublished doct diss University of Wisconsin-Madison, 1994.
37 Silverman K: The Structure and Processing of Fundamental Frequency Contours; unpublished doct diss University of Cambridge, 1987.
38 Papso C, Blood I: Word recognition skills of children and adults in background noise. Ear Hear 1989; 10:235–236.
39 Sperry J, Wiley T, Chial M: Word recognition performance in various background competitors. J Am Acad of Audiol 1997;8:71–80.
40 Danhauer J, Leppler J: Effects of four noise competitors on the California consonant test. J Speech Hear Disord 1979;44:354–362.
41 Schum DJ: Speech understanding in background noise; in Valente M (ed): Hearing Aids: Standards, Options, and Limitations. New York, Thieme, 1996, pp 368–406.
42 Weismer G, Jeng J, Laures J, Kent R, Kent J: Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders. Folia Phoniatr Logop 2001;53:1–18.
43 Slowiaczek L, Nusbaum H: Effects of speech rate and pitch contour on the perception of synthetic speech. Hum Factors 1985;27:701–712.
44 Nusbaum H, Pisoni D: Constraints on the perception of synthetic speech generated by rule. Behav Res Methods Instrum Comput 1985;17:235–242.
45 Parkhurst B, Levitt H: The effect of selected prosodic errors on the intelligibility of deaf speech. J Commun Disord 1978;11:249–256.
46 Stevens K, Nickerson R, Rollins A: Suprasegmental and postural aspects of speech production and their effect on articulatory skills and intelligibility; in Hochberg I, Levitt J, Osberger M (eds): Speech of the Hearing-Impaired Research Training and Personnel Preparation. Baltimore, University Park Press, 1983, pp 33–51.
47 Kewley-Port D, Li X, Zheng Y, Neel A: Fundamental frequency effects on thresholds for vowel formant discrimination. J Acoust Soc Am 1992;100:2462–2470.
48 Bencala K: The interaction effects of vocal tract scaling (male-to-female) and fundamental frequency on vowel identification; unpublished master's thesis University of Arizona, 2003.
49 Bunton K: An acoustic item-analysis of intelligibility; unpublished doct diss University of Wisconsin-Madison, 1999.
50 Yorkston K, Buekelman D, Strand E, Bell K: Management of Motor Speech Disorders in Children and Adults. Austin, Pro-Ed, 1999.
51 Scott S, Caird F: The response of the apparent perceptive speech disorder of Parkinson's disease to speech therapy. J Neurol Neurosurg Psychiatry 1984;47:302–304.
52 Scott S, Caird F: Speech therapy for patients with Parkinson's disease. J Neurol Neurosurg Psychiatry 1983;46:140–144.
53 Scott S, Caird F: Speech therapy for patients with Parkinson's disease. BMJ 1981;283:168.
54 Ramig L, Countryman S, O'Brien C, Hoehn M, Thompson L: Intensive speech treatment for patients with Parkinson's disease: short- and long-term comparison of two techniques. Neurology 1996;47: 1496–1504.
55 Ramig L, Sapir S, Fox C, Countryman S: Changes in vocal loudness following intensive voice treatment (LSVT) in individuals with Parkinson's disease: a comparison with untreated patients and normal age-matched controls. Mov Disord 2000;16:79–83.