0743-4618/01/1801-0002 $3.00/0; Volume 18, March 2002
AAC Augmentative and Alternative Communication

Phonatory Control in Adults with Cerebral Palsy and Severe Dysarthria

Rupal Patel
Department of Biobehavioral Studies, Teachers College, Columbia University, New York, New York; Speech Communication Group and The Media Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

Augmentative and alternative communication (AAC) users with severe dysarthria may benefit greatly from using residual vocalizations to enhance communication efficiency and naturalness. Many individuals use their residual vocal control to gain attention, express emotions, and convey intentions to familiar care givers. This research was designed to test the hypothesis that speakers signal such information through the control of phonatory features such as pitch and duration, despite severe degradation of segmental speech sounds. The study examined whether eight speakers with severe dysarthria caused by cerebral palsy could control the pitch and duration of sustained vowel productions at three distinct levels. The results indicated that all speakers were able to consistently control sustained production of the vowel /a/ at three durations (short, medium, and long). Speakers were more variable, however, in their ability to control pitches (low, medium, and high). Seven of the eight speakers were able to produce at least two distinct pitches. The potential impact of harnessing residual phonatory control as an additional channel of input for AAC users with severe speech and motor impairment is discussed.

KEY WORDS: augmentative and alternative communication (AAC), cerebral palsy, communication bandwidth, dysarthria, input, prosody, voice-driven interface

Individuals with severe dysarthria caused by cerebral palsy are unable to communicate using speech alone. They typically compensate for their disorder by using augmentative and alternative communication (AAC) systems. Accessing a communication aid by pointing and/or scanning is slow, cumbersome, and unnatural and can be physically fatiguing (Ferrier, Jarrell, Carpenter, & Shane, 1992; Treviranus, Shein, Haataja, Parnes, & Milner, 1991; Vanderheiden, 1985). Markedly reduced rates of communication imposed by these modes of access (Mathy-Laikko, West, & Jones, 1993; Vanderheiden, 1985) constrain the efficiency and dynamics of communication exchange between AAC users and their communication partners (Beukelman & Mirenda, 1998; Shein, Brownlow, Treviranus, & Parnes, 1990).

Communication rate may be improved by enabling AAC users to convey information via several modes including pointing, scanning, and voice. These access modes may be used either simultaneously to signal various aspects of a message or in a redundant fashion to cope with fatigue in any one modality. It has been documented that many AAC users effectively use multiple input modes with familiar communication partners (Allaire, Gressard, Blackman, & Hostler, 1991; Rounsfell, Zucker, & Roberts, 1993; Shein et al., 1990; Smith, 1994). Most currently available AAC devices, however, are designed to attend to the user's most reliable and consistent mode of input.

The use of vocalizations as an input mode has the potential to improve communication efficiency. Treviranus et al. (1991) reported increased rates of input using scanning and voice as a hybrid input method when compared with scanning alone. In addition, some AAC users may find that using vocalizations is a more natural interface for communication than using nonvocal input, which tends to alter the social dynamics of conversation (Beukelman & Mirenda, 1998).

For individuals with severe dysarthria, vocalization access is rarely considered, given the nature and degree of their speech impairment. It is not uncommon, however, to find AAC users who exploit all available vocalizations when interacting with familiar communication partners (e.g., Allaire et al., 1991; Ferrier, 1991; Ferrier, Shane, Ballard, Carpenter, & Benoit, 1995; Fried-Oken, 1985; Murphy, Markova, Moodie, Scott, & Boa, 1995; Smith, 1994).

How do familiar communication partners decipher the intended meaning of dysarthric utterances? Perhaps situational cues (e.g., facial expressions, body language, emotional state), familiarity and experience with the speaker (see McGarr, 1983, and Monsen, 1983, for analogous studies with hearing-impaired populations), and linguistic context (Beliveau, Hodge, &

Hagler, 1995; Dongilli, 1993; Yorkston, Hammen, & Dowden, 1991) provide communication partners with cues regarding the intended message. In addition to these contextual cues, there may also be consistencies in the acoustic features of dysarthric vocalizations.

Characterization of dysarthria caused by cerebral palsy has typically focused on the phonetic aspects of the speech signal (Ansel & Kent, 1992; Platt, Andrews, & Howie, 1980; Platt, Andrews, Young, & Quinn, 1980). Imprecise consonants, distorted vowels, and hypernasality are commonly cited perceptual characteristics (Darley, Aronson, & Brown, 1969, 1975; Rosenbek & LaPointe, 1978; Yorkston, Beukelman, Strand, & Bell, 1999). Associated acoustic features include (a) centralized vowel quadrangles, (b) increased bandwidths of the first and second formants, (c) increased interformant noise, (d) difficulty with formant transitions, (e) loss of acoustic contrasts, (f) difficulty with fricative and affricate sounds, (g) weaker plosives, (h) unexpected nasalization, and (i) high intra- and interspeaker variability (Ansel & Kent, 1992; Doyle, Raade, St. Pierre, & Desai, 1995; Jayaram & Abdelhamied, 1995; Platt, Andrews, & Howie, 1980; Platt, Andrews, Young, et al., 1980; Rosenbek & LaPointe, 1978).

Speech sound intelligibility may also be affected by phonatory features such as pitch and duration (Beukelman & Yorkston, 1979; Bunton, Weismer, & Kent, 2000; Laures & Weismer, 1999; Liss, Spitzer, Caviness, Alder, & Edwards, 1998; Traunmuller, 1981; Yorkston et al., 1999). Dysarthric vocalizations have been described as monopitch, monoloud, and slow in rate (Darley et al., 1969, 1975; Hardy, 1983; Le Dorze, Ouellet, & Ryalls, 1994; Wit, Maassen, Gabreels, & Thoonen, 1993). Empirical studies also document reduced range in fundamental frequency (F_0) for dysarthric speakers compared to normal controls (Canter, 1963; Darley et al., 1975; Jacques, Rastatter, & Sullivan, 1985; Kent & Rosenbek, 1982; Schlenck, Bettrich, & Willmes, 1993; Wit et al., 1993). Similar reductions in range have also been noted for vowel prolongation (Darley et al., 1975; Hardy, 1983). In addition to narrowed range, greater variability on repeated productions has also been reported among individuals with dysarthria compared to normal speakers (Darley et al., 1969; Hardy, 1983; Platt, Andrews, & Howie, 1980; Platt, Andrews, Young, et al., 1980).

It is apparent from the previous section that dysarthric abilities are typically expressed in the literature in terms of deviations from normal speech. The labels used tend to focus on the reduced precision and flexibility of the dysarthric vocal apparatus in comparison to normal. As Abberton, Fourcin, and Hazan (1991) noted, however, F_0 range and the use of pitch within that range are separate issues. Furthermore, cataloguing differences between normal and dysarthric control does not help explain the nature and degree of control available within the narrowed range or how to harness any residual control.

Furthermore, various subsystems of speech production—respiratory, laryngeal, and articulatory—may be impaired to differing levels of severity within individuals with dysarthria caused by cerebral palsy (Blumberg, 1955; Hardy, 1983). Thus, it is possible that individuals with relatively spared respiratory and laryngeal function but poor articulatory control are better able to manipulate phonatory features when compared to speech sound precision.

It may also be possible to learn more about the compensatory strategies employed by individuals with dysarthria by studying their phonatory abilities. Speakers may remap some aspects of vocal control to features that are easier for them to produce (Brewster, 1989; Crystal, 1979; Vance, 1994). For example, Vance (1994) reported a case study of a 30-year-old male with mild to moderate dysarthria who consistently used low falling pitch to substitute for rising pitch because of ease of production.

It may also be the case that speakers with markedly reduced control over phonetic aspects of the speech signal still have sufficient control to consistently convey their intentions by varying the pitch and duration of vocalizations (Patel, 1998; Patel & Roy, 1998; see also Le Dorze et al., 1994; Vance, 1994; Wit et al., 1993). These phonatory parameters span a longer temporal window than more transient units such as phonemes. Given that dysarthria is a motor speech impairment often characterized by slow, weak, and imprecise movement, it follows that transient articulatory gestures would be more difficult to control than the relatively slow and gradually varying features of pitch and duration.

This study investigated the ability of individuals with severe dysarthria caused by cerebral palsy to control pitch and duration during sustained vowel production. Speakers were explicitly instructed to control pitch or duration of a vowel at three distinct levels for each (low, medium, and high for pitch; short, medium, and long for duration). Each speaker's ability to control pitch and duration was considered separately. The purpose of the study was to determine whether speakers could produce three distinct and separable categories within each parameter. In contrast to previous studies, the goal was to study the degree and nature of phonatory control available to individuals with severe dysarthria rather than to compare these individuals to normal control subjects.

METHOD

Participants

A group of eight speakers with severe dysarthria resulting from cerebral palsy participated in this experiment. Speakers ranged in age from 27 to 44 years, with a mean age of 36.4 years. Speakers met a set of four selection criteria. First, dysarthria was the primary speech diagnosis for all participants. The investigator performed a battery of formal and informal evaluations to determine the nature and degree of speech impairment. This consisted of screening for adequate receptive language skills, an oral peripheral examination, a differential diagnosis of the dysarthria, a thorough case history, and an interview to gather information about present and preferred modes of communication. Second, the severity of the participants' dysarthria made it necessary for them to use AAC for daily communication. A modified version of the *Assessment of Intelligibility of Dysarthric Speakers* (Yorkston & Beukelman, 1981) was administered to determine the level of severity of dysarthria. To minimize speaker fatigue, 25 rather than 50 isolated word productions were requested. The average of the two unfamiliar raters' scores was used as a measure of the speakers' intelligibility. Third, all speakers were required to pass an audiometric evaluation. Fourth, grossly adequate cognitive skills were necessary for participation. Table 1 summarizes the participant characteristics.

Setting

To ensure controlled experimental conditions, collection of speech samples took place in a sound-treated, comfortable, adult-appropriate, wheelchair-accessible audiometric booth.

Materials

Experimental protocols were audiotaped using the Sony® Digital Audiorecorder (Model PCM-2300) on 90-minute Sony digital audiotapes. The Shure® professional unidirectional head-mount cardioid dynamic microphone (Model SM10A) was placed on the speaker and adjusted at a distance of 2 cm from the left corner of the mouth. A constant mouth-to-microphone distance was maintained and monitored carefully throughout the session.

The Praat system,[1] a general-purpose speech analysis package, was used to analyze vocalizations. Statistical analyses were performed using SPSS (Version 9.0).[2]

Design

This experiment followed a quasiexperimental design. Control of pitch and duration for sustained production of the vowel /a/ was considered separately using two identical protocols that were both collected on the same day. The order of collection of pitch and duration protocols was counterbalanced among all eight speakers. Within each protocol, requests for distinct pitches and durations were randomized to minimize confounding practice effects and/or fatigue effects.

Dependent Measures

For the pitch protocol, the dependent measure was the average F_0 of each vocalization measured in Hertz. In the duration protocol, the dependent measure was the length of each vocalization measured in seconds.

Procedure

In each protocol, speakers were asked to produce 20 vocalizations[3] at three distinct levels. In the pitch protocol, the three levels corresponded to low-, medium-, and high-pitched productions. In the duration protocol, the three levels corresponded to short, medium, and long productions.

Initial Establishment Phase

At the start of each protocol, the experimenter produced examples of vocalizations to demonstrate the task. Each participant had the opportunity to practice producing vocalizations at different pitches or durations. Next, participants established their pitch or duration range. In this initial establishment phase, participants were able to monitor their productions using Visi-Pitch,[4] a graphical display of frequency versus time. Visual feedback from Visi-Pitch was helpful in illustrating the concepts of the experimental task. Once a range was established, the speaker was asked to produce a vocalization that he/she judged to be high, medium, and low within that range. Visual stickers were affixed to the Visi-Pitch display to provide guides for the different levels. Participants were allowed five trials to practice replicating each level. The Visi-Pitch was then removed, and participants were asked to perform three additional practice trials of each level (i.e., high, medium, and low) without visual feedback.

Data Collection Phase

During the data collection phase, visual feedback using the Visi-Pitch was not provided. In the pitch pro-

---

[1] Boersma P, Weenink D. (1990). Praat: A System for Doing Phonetics by Computer (Version 3.80) [shareware]. Amsterdam, Neimejen, Netherlands: Institute of Phonetics Sciences of the University of Amsterdam.

[2] SPSS Inc. Headquarters, 233 S. Wacker Drive, 11th Floor, Chicago, IL, 60606, USA.

[3] To avoid small sample bias, a very large sample size would be needed for each level of the prosodic parameter. This was not feasible, however, given that vocal fatigue was a serious concern. Selecting a total sample size of 20 trials was a compromise. Even if vocalizations within each level of the prosodic parameter were not normally distributed, the total "population" of vocalizations would approximate a normal distribution for a total sample size of 60 trials (Pagano & Gauvreau, 1993).

[4] Visi-Pitch Model 6087 DS, Kay Elemetrics Corporation, 12 Maple Ave, Pine Brook, NJ 07058, USA.

TABLE 1: **Description of Speakers with Dysarthria**

Speaker	*Age*	*Gender*	*Speech Intelligibility (%)*	*Reported Modes of Communication*	*Motor Control*
S1	27	M	12	Gesture, vocalizations, sign language	Able to sign, unable to write, uses a wheelchair
S2	33	F	18	Head pointer, vocalizations	Uses a wheelchair, head control
S3	35	M	20	Communication board with Blissymbols, vocalizations	Uses a wheelchair, pointing gestures
S4	32	M	22	Picture symbols, some vocalizations	Uses a wheelchair, pointing gestures
S5	44	F	16	Alphanumeric communication board, some vocalizations	Uses a wheelchair, pointing gestures
S6	36	M	22	Alphanumeric communication board, some vocalizations	Ambulatory, pointing gestures
S7	40	M	18	Alphanumeric and phrase communication board	Uses a wheelchair, pointing gestures
S8	44	M	24	Alphanumeric board, vocalizations with familiar people	Ambulatory, pointing gestures

tocol, speakers were asked to produce 20 vocalizations at each pitch (i.e., high, medium, and low). In the duration protocol, speakers produced 20 vocalizations at each duration (i.e., short, medium, and long). Approximately 3 seconds were allowed between the end of the participant's last production and the request for the next vocalization. If a participant produced a vocalization and then spontaneously corrected himself/herself, the self-correction was accepted as the actual trial, though only one reattempt was allowed. Spontaneous self-corrections were assumed to be accidental errors caused by a lack of concentration or anticipation of the next trial rather than a lack of control. A rest period of at least 10 minutes was provided between protocols.

Preparation of Data for Acoustic Analysis

For each participant, the 60 vocalizations collected from each of the two protocols were analyzed. The Praat System was used to visualize the productions and to calculate the average pitch and duration values for each vocalization. First, the beginning and end of each vocalization trial were manually marked. Intrajudge reliability for manual placement of cursor positions was calculated for a random sample of 10% of the data. The Pearson correlation coefficient between labeling sessions was 0.9962. A paired *t*-test between the original cursor placements and repeated markings was nonsignificant ($p = .239$). For the duration protocol, the duration in seconds of the marked region of each vocalization was measured. In the pitch protocol, the average F_0 of each vocalization (Hz) was calculated.

RESULTS

Each speaker's ability to control pitch and duration was considered separately. The data were analyzed to determine whether speakers could produce three distinct categories in each parameter.

Control of Pitch

Separate analyses of variance were completed for each speaker and each protocol using SPSS. In the pitch protocol, the analysis was carried out on the 60 data values of average F_0. The mean and standard deviation of the low, medium, and high distributions were computed using the corresponding 20 data points for each requested level of F_0. Figure 1 shows the results for all speakers. These values were used to form histograms for each level to compare the degree of overlap between adjacent categories.

For any given speaker, the distributions of low, medium, and high pitch were not symmetrically distributed about the corresponding means and the variances of the distributions were not equal. Standard transformations, including data inversion, square roots, inverse square roots, and natural logarithms, were applied; however, the transformed data still violated assumptions of parametric analysis for some speakers. For this reason, a nonparametric approach was taken using the Kruskal-Wallis H test. This nonparametric statistic does not assume normal underlying distributions and is used to determine whether a group of independent samples is from the same or different populations (Downie & Heath, 1965). Statistically significant main effects between high, medium, and low pitch sounds

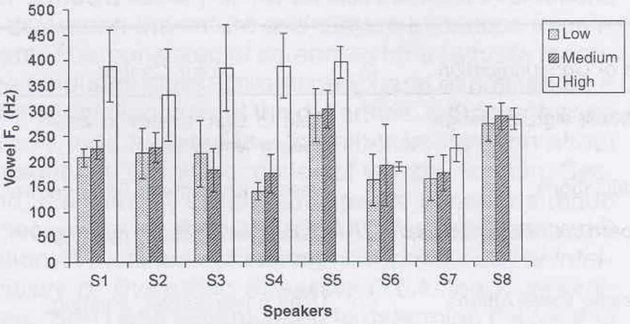

Figure 1. Mean F_0 and standard deviation per pitch category (low, medium, and high) for all speakers.

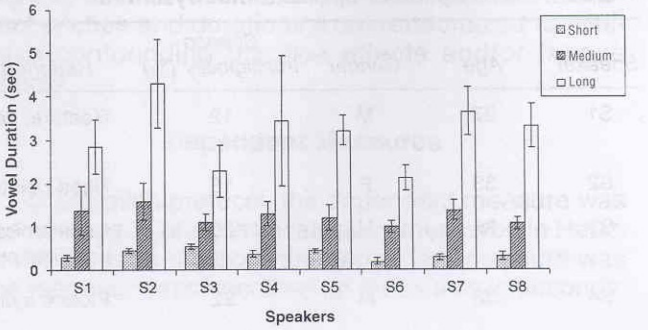

Figure 2. Mean duration and standard deviation per duration level (short, medium, and long) for all speakers.

were found for all but one speaker ($p < .05$). Speaker S8 did not have significant differences in F_0 for high, medium, or low categories ($p = .0301$).

To determine which pitch levels differed significantly from each other, Tamhane's T2 test was performed. This a posteriori test performs conservative pairwise comparisons and is appropriate when the variances between comparison groups cannot be assumed to be equal. The Tamhane procedure uses the Student's *t* distribution and controls for the experiment-wise error rate at or below the predetermined α level (Kirk, 1982). Table 2 summarizes the results of multiple comparisons between the three levels of pitch requested. The exact p values are listed in the table, and significant differences are marked with an asterisk.

Speakers S4 and S1 were able to control all levels of pitch, resulting in a significantly different F_0 for each target pitch level ($p < .05$). Speaker S2 was able to produce high pitch sounds that significantly differed from low and medium pitch levels. The distributions of low and medium pitch sounds, however, were not significantly different ($p = .385$). Similar findings were also true for speakers S3, S5, and S7, who were all able to produce significantly different high versus low and high versus medium pitches but were unable to differentiate between low and medium pitches. Speaker S6 was able to produce significantly different low and high pitches ($p < .05$), but neither distribution differed significantly from the medium pitch level.

Control of Duration

Distributions of vocalization duration were determined using the 20 data points collected for each target level. Figure 2 illustrates the results of the duration protocol.

Compared to the pitch protocol, the distributions of duration were relatively Gaussian (i.e., followed a normal distribution). Long duration, however, had a much higher variance than the short- or middle-duration distributions for almost all speakers. This violation of the assumption of equal variance again necessitated the use of the Kruskal-Wallis H test. The results indicated that all speakers were able to produce at least two different durations ($p < .05$).

To determine which durations differed significantly from each other, Tamhane's T2 test was performed. All speakers were able to produce three distinct durations, with all contrasts reaching significance ($p < .0001$). Figure 3 provides a summary of the number of estimated categories of pitch and duration for each speaker.

TABLE 2: **Summary of Statistically Significant Contrasts Between Low, Medium, and High F_0 Levels**

Speaker	*Contrast*	*Probability*
S1	Low vs. medium	.017*
	Medium vs. high	.000*
	Low vs. high	.000*
S2	Low vs. medium	.385
	Medium vs. high	.000*
	Low vs. high	.010*
S3	Low vs. medium	.214
	Medium vs. high	.000*
	Low vs. high	.000*
S4	Low vs. medium	.001*
	Medium vs. high	.000*
	Low vs. high	.000*
S5	Low vs. medium	.794
	Medium vs. high	.000*
	Low vs. high	.000*
S6	Low vs. medium	.418
	Medium vs. high	.921
	Low vs. high	.027*
S7	Low vs. medium	.570
	Medium vs. high	.000*
	Low vs. high	.000*
S8	Low vs. medium	.313
	Medium vs. high	.863
	Low vs. high	.603

*Statistically significant contrast.

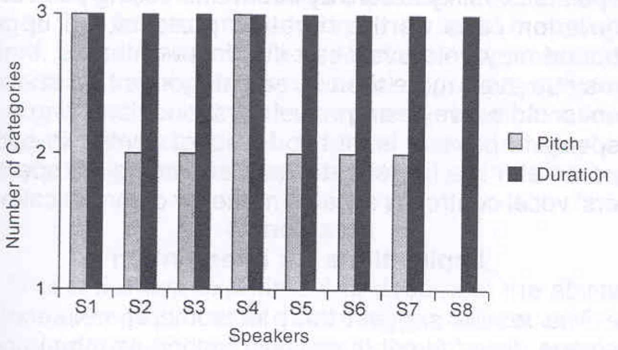

Figure 3. Number of estimated categories of pitch and duration for each speaker.

DISCUSSION

Control of Pitch

The pitch protocol was used to determine the number of separate and distinct levels of F_0 for each speaker. Speakers differed markedly in their ability to control F_0 for sustained production of the vowel /a/. Only two speakers, S1 and S4, were able to produce three minimally overlapping, distinct categories of F_0. The remaining speakers had two levels of control (S2, S3, S5, S6, and S7); speaker S8 had no detectable control.

Mixed results across speakers may indicate differences in the degree of laryngeal involvement of each speaker. Speakers who were able to produce two or more categories of F_0 may have had less laryngeal involvement than speaker S8, who was unable to perform the task.

The overall range of F_0 varied considerably across speakers. Speaker S8 had an extremely reduced range, with productions for all three pitches falling between 180 and 380 Hz. Thus, speaker S8 was unable to produce reliable pitch differences. Speaker S7 also had a narrow range of F_0, from 95 to 305 Hz. He was, however, able to signal the difference between low and high pitch productions. Speakers S2, S3, S5, and S6 had a considerably larger F_0 range, in some instances from 95 to 515 Hz. Consequently, these speakers were able to signal at least two distinct pitch categories. Although speakers S1 and S4 did not have as great an F_0 range as S2 or S3, they were able to exert finer control to produce three distinct distributions. Although overall F_0 range seems to be correlated with discriminability into categories, it is not the only factor influencing the number of distinct pitch levels.

Inconsistency of productions across trials within a given pitch level was also evident. Distributions for each level requested were rarely unimodal. Unimodal distributions would indicate that the speaker's productions were clustering about a given mean with some allowable variation between trials. On the other hand, bimodal or uniform distributions would indicate that the speaker was not able to consistently achieve a single target F_0 level. Some speakers were able to produce some pitch levels with more consistency than other pitch levels. For example, speaker S4 was able to produce highly consistent productions for low and medium pitch, but his high pitch productions spanned a larger F_0 range and the distribution was relatively uniform. In contrast, speaker S6 had more consistency in producing high pitch but less consistency with both low and medium pitch levels. Many speakers had F_0 distributions with two or more modes. A trend toward greater variance and a non-normal distribution of F_0 for the high pitch level was evident for four (S1, S2, S3, and S4) of eight speakers.

Some of the within-speaker inconsistencies may have been attributable to the small sample size. With more trials, the distributions may have tended toward normal distributions, but the practical difficulties of collecting larger sample sizes restricted the data set. However, sampling size may not be the only relevant factor; inconsistency in repeated productions may also be attributable to an inherent lack of control.

The findings of inconsistent production within a given pitch level may also have resulted from the operational definition of pitch control used in this experiment. For each trial, the average F_0 of the vowel was used as the pitch measure. Examination of the productions for many speakers showed a great deal of F_0 variability within the vowel prolongation task. That is, the vowel was not steadily maintained at a specific F_0 level throughout the production. In many cases, there was a drift in F_0 toward the target pitch throughout the progression of the vocalization. In these instances, a measure of the final few milliseconds of the production may have been more useful. In other cases, however, speakers would start off at a lower level, attempt to hit the target level, and then would drop the pitch at the end of the utterance. In these instances, a measure of the peak F_0 may have been a better indicator of pitch control.

Control of Duration

The duration protocol was used to determine the number of separate and distinct levels of vowel duration for each speaker. The results indicated that all eight speakers were able to produce vowels at three distinct, minimally overlapping levels. For almost all speakers, except perhaps S3, visual inspection of the distance between distributions of productions indicated that they might have been able to control more than three levels of duration. In many cases, adjacent distributions did not overlap at all. In such cases, it is conceivable that speakers may be able to produce an additional level of vowel duration. For example, if the short and medium durations were nonoverlapping and well separated, a speaker may have been able to produce vocalizations of "short-medium" duration, a category between short and medium, if it had been requested. There is, however, a limit to the number of categories along a given continuum. The greater the number of categories requested, the poorer the separation between adjacent categories (Garner, 1953). It is important to be mindful of the trade-off between the number of categories requested and the potential degree of overlap between adjacent categories if this control is used to operate a communication device. The greater the overlap between categories, the higher the probability for error.

In contrast to the pitch protocol, the distributions for the duration protocol tended to be relatively unimodal for all three levels requested across all eight speakers. Even speaker S8, who was unable to control pitch, was able to produce vocalizations that varied in duration.

As with the pitch protocol, the long-duration category tended to have greater variance and range than the short-duration category. This tighter clustering for the short-duration sounds may be attributable, at least partially, to edge effects. Categories at the edge of a continuum tend to be more tightly clustered than categories in the middle, given that edge categories often serve as perceptual anchors (Allen, 1994; Garner, 1953; Garner & Hake, 1951; Miller, 1954). Although this theory would also predict edge effects in the long-duration category, this prediction was not supported by the results of this study. Although speakers may be able to produce long vowel sounds that differ from short- and medium-duration sounds, insufficient respiratory support (Blumberg, 1955; Hardy, 1983) may cause greater variance in the length of long vowels. Vocal fatigue may also have contributed to greater variability for long productions than for short and medium productions.

Similar to the results reported by Wit et al. (1993), the results of the present study indicated a reduced range in duration. For example, speakers were able to prolong the vowel /a/ for up to 6 seconds, only a third of the duration typical of normal speakers (Finnegan, 1984). Despite this narrowed range, the degree of utterance-to-utterance variability was sufficiently small such that all eight speakers were able to produce at least three distinct vowel durations. This finding is contrary to previous accounts of increased within-speaker variability (Darley et al., 1969; Hardy, 1983; Platt, Andrews, & Howie, 1980; Platt, Andrews, Young, et al., 1980). Had the present analysis been limited to comparisons of normal and dysarthric speakers with regard to their ability to control vowel duration, the fact that the dysarthric participants were still able to exert functional control over duration within a narrowed range would not have been detected.

To summarize, this research examined the vocal abilities of individuals with severe dysarthria within a framework of understanding the information transmission capacity in each prosodic parameter. The focus was to determine an upper bound on each speaker's ability to convey information using pitch and duration cues. In the duration protocol, an upper bound may not have been determined; that is, there may be even more than three categories that speakers could convey using vowel duration. Identifying that speakers have at least two levels of control in each parameter is a first step toward leveraging the speakers' vocal control as a viable mode for communication.

Implications for Intervention

The results suggest that, for some speakers with severe dysarthria, it is possible to increase the bandwidth of information transmitted if they are instructed to produce vowels of contrastive pitches and/or durations. Interventions may be designed to optimize a speaker's expressive communication abilities by training the individual to produce consistently distinct duration and F_0 cues to signal various messages. A speaker's current vocalization repertoire may be shaped to produce an increased number of distinguishable utterances by superimposing patterns of pitch or duration modulation on a finite set of consistent articulatory gestures (i.e., perhaps the cardinal vowels). This method of deliberately using residual pitch and duration control to increase the number of communicative messages has not been used to date with speakers with severe dysarthria as a mode of intervention.

In addition to shaping the vocal behaviors of AAC users with dysarthria, it may also be possible to train communication partners to be more attuned to subtle pitch and duration cues that are communicative. Communication partners may include human listeners and also dedicated computer systems that are capable of discriminating among vocalizations based on phonatory cues in the vocalization signal. In the future, such treatment/training regimens may afford some AAC users with an additional channel of input to their voice output communication aid. Using voice may reduce physical fatigue, improve efficiency, and improve the naturalness of interaction for many AAC users with severe speech and motor impairment.

Future Directions

Further analyses on the current data set are under way to search for an optimal measure of F_0 that would allow for maximal separation of the three pitch categories. For example, alternative measures such as slope of pitch, rather than average F_0, may improve the results. Future research plans include collection of a larger corpus of data from speakers on multiple occasions to further investigate the consistency of their vocal productions. Additional research will also aim to determine whether speakers can vary pitch and duration simultaneously. If so, the number of possible combinations of vocal control could be multiplied.

Although it may be possible to teach an individual to map consistent prosodic categories of vowel productions to a set of messages, this technique is limiting in that the user must learn a new input mapping. Thus, future research will also include experiments to determine the phrase level control of prosody, which, if consistent, could further improve communication efficiency and may be more natural than learning arbitrary vocalizations to map meanings.

Conclusion

This is the first experiment to document the ability of speakers with severe dysarthria to produce vowels at contrastive pitches and durations. All eight speakers with severe dysarthria caused by cerebral palsy were able to produce at least three distinct durations consistently. Speakers were more variable in their ability to control pitch; however, all but one speaker successfully marked the difference between low and high pitches. At the very least, the speakers were capable of producing a basic binary contrast in vowel pitch. Using a simple Morse code system of distinct durations and/or pitches, a speaker with severe dysarthria may be able to transmit a considerable amount of information using his/her voice. This potential to harness phonatory control for conveying information has tremendous implications for improving communication efficiency and naturalness.

ACKNOWLEDGMENTS

This research was conducted in the Department of Speech-Language Pathology at the University of Toronto toward partial fulfillment of the requirements for a doctoral degree. I am grateful to Bernard O'Keefe for his mentorship, support, and guidance. I would also like to thank Hans Kunov, Christopher Dromey, Luc DeNil, and Deb Roy for their contributions to this work. This study would not have been possible without the involvement of the participants and the support of the National Science and Engineering Research Council of Canada, Bloorview Children's Hospital Foundation, Whipper Watson Foundation, Donalda MacGeachy Foundation, and the Harmonize of Speech Foundation of Ontario.

Address reprint requests to: Rupal Patel, Department of Biobehavioral Studies, Teachers College, Columbia University, 525 W. 120th Street, New York, NY, 10027, USA; e-mail: patel@exchange.tc.columbia.edu.

REFERENCES

Abberton, E., Fourcin, A., & Hazan, V. (1991). Fundamental frequency range and the development of intonation in a group of profoundly deaf children. *Proceedings of the XII International Congress of Phonetic Sciences, 5*, 142–145.

Allaire, J. H., Gressard, R. P., Blackman, J. A., & Hostler, S. L. (1991). Children with severe speech impairment: Caregiver survey of AAC use. *Augmentative and Alternative Communication, 7*, 248–255.

Allen, J. B. (1994). How do humans process and recognize speech? *IEEE Transactions on Speech and Audio Processing, 2*, 567–577.

Ansel, B. M., & Kent, R. D. (1992). Acoustic-phonetic contrasts and intelligibility in the dysarthria associated with mixed cerebral palsy. *Journal of Speech and Hearing Research, 35*, 296–308.

Beliveau, C., Hodge, M. M., & Hagler, P. H. (1995). Effects of supplemental linguistic cues on the intelligibility of severely dysarthric speakers. *Augmentative and Alternative Communication, 11*, 176–186.

Beukelman, D. R., & Mirenda, P. (1998). *Augmentative and alternative communication: Management of severe communication disorders in children and adults* (2nd ed.). Baltimore: Paul H. Brookes.

Beukelman, D. R., & Yorkston, K. M. (1979). The relationship between information transfer and speech intelligibility of dysarthric speakers. *Journal of Communication Disorders, 12*, 189–196.

Blumberg, M. (1955). Respiration and speech in the cerebral palsied child. *American Journal of Disorders of Childhood, 89*, 48–53.

Brewster, K. (1989). The assessment of prosody. In K. Grundy (Ed.), *Linguistics in clinical practice* (pp. 186–202). London: Whurr.

Bunton, K., Weismer, G., & Kent, R. (2000, February). *The effects of flattened fundamental frequency contours on sentence intelligibility in dysarthric speakers.* Paper presented at the conference on motor speech disorders and speech motor control, San Antonio, TX.

Canter, G. J. (1963). Speech characteristics of patients with Parkinson's disease: Intensity, pitch and duration. *Journal of Speech and Hearing Disorders, 28*, 221–228.

Crystal, D. (1979). Prosodic development. In P. Fletcher & M. Garman (Eds.), *Language acquisition* (pp. 174–197). Cambridge, UK: Cambridge University Press.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969). Differential diagnostic patterns of dysarthria. *Journal of Speech and Hearing Research, 12*, 246–269.

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

Dongilli, P. A. Jr. (1993). *Semantic context and speech intelligibility.* Unpublished doctoral dissertation, University of Nebraska-Lincoln.

Downie, N. M., & Heath, R. W. (1965). *Basic statistical methods.* New York: Harper & Row.

Doyle, P. C., Raade, A. S., St. Pierre, A., & Desai, S. (1995). Fundamental frequency and acoustic variability associated with production of sustained vowels by speakers with hypokinetic dysarthria. *Journal of Medical Speech-Language Pathology, 3*, 41–50.

Ferrier, L. J. (1991). Clinical study of a dysarthric adult using a Touch Talker with Words Strategy. *Augmentative and Alternative Communication, 7*, 266–274.

Ferrier, L. J., Jarrell, N., Carpenter, T., & Shane, H. (1992). A case study of a dysarthric speaker using the DragonDictate voice recognition system. *Journal for Computer Users in Speech and Hearing, 8*, 33–52.

Ferrier, L. J., Shane, H. C., Ballard, H. F., Carpenter, T., & Benoit, A. (1995). Dysarthric speakers' intelligibility and speech characteristics in relation to computer speech recognition. *Augmentative and Alternative Communication, 11*, 165–173.

Finnegan, D. E. (1984). Maximum phonation time for children with normal voices. *Journal of Communication Disorders, 17*, 309–317.

Fried-Oken, M. (1985). Voice recognition device as a computer interface for motor and speech impaired people. *Archives of Physical Medicine and Rehabilitation, 66*, 678–681.

Garner, W. R. (1953). An informational analysis of absolute judgements of loudness. *Journal of Experimental Psychology, 46*, 373–380.

Garner, W. R., & Hake, H. W. (1951). The amount of information in absolute judgements. *Psychological Review, 58,* 446–459.

Hardy, J. C. (1983). *Cerebral palsy.* Englewood Cliffs, NJ: Prentice-Hall.

Jacques, R. D., Rastatter, M., & Sullivan, J. (1985). Some effects of congenital spasticity on fundamental frequency. *Perceptual Motor Skills, 61,* 75–80.

Jayaram, G., & Abdelhamied, K. (1995). Experiments in dysarthric speech recognition using artificial neural networks. *Journal of Rehabilitation Research and Development, 32,* 162–169.

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

Kirk, R. E. (1982). *Experimental design: Procedures for the behavioral sciences.* Belmont, CA: Brooks/Cole.

Laures, J. S., & Weismer, G. (1999). The effects of a flattened fundamental frequency contour on intelligibility at the sentence level. *Journal of Speech, Language and Hearing Research, 42,* 1148–1156.

Le Dorze, G., Ouellet, L., & Ryalls, J. (1994). Intonation and speech rate in dysarthric speech. *Journal of Communication Disorders, 27,* 1–18.

Liss, J. M., Spitzer, S., Caviness, J. N., Alder, C., & Edwards, B. (1998). Syllabic strength and lexical boundary decisions in the perception of hypokinetic dysarthric speech. *Journal of the Acoustical Society of America, 104,* 2457–2466.

Mathy-Laikko, P., West, C., & Jones, R. (1993). Development and assessment of a rate acceleration keyboard for direct selection augmentative and alternative communication users. *Technology and Disability, 2,* 57–67.

McGarr, N. S. (1983). The intelligibility of deaf speech to experienced and inexperienced listeners. *Journal of Speech and Hearing Research, 26,* 451–458.

Miller, G. A. (1954). Note on the bias of information estimates. In H. Quastler (Ed.), *Information theory and psychology: Problems and methods* (pp. 95–100). Glencoe, IL: Free Press.

Monsen, R. B. (1983). The oral speech intelligibility of hearing impaired talkers. *Journal of Speech and Hearing Disorders, 48,* 286–296.

Murphy, J., Markova, I., Moodie, E., Scott, J., & Boa, S. (1995). AAC systems used by people with cerebral palsy in Scotland: A demographic study. *Augmentative and Alternative Communication, 11,* 26–36.

Pagano, M., & Gauvreau, K. (1993). *Principles of biostatistics.* Belmont, CA: Duxbury Press.

Patel, R. (1998). Acoustic analysis of vocalizations produced by individuals with severe dysarthria. In *Proceedings of the International Society of Alternative and Augmentative Communication Biennial Conference* (pp. 489–491). Dublin: ISAAC.

Patel, R., & Roy, D. (1998). Teachable interfaces for individuals with dysarthric speech and severe physical disabilities. In *Proceedings of the AAAI Workshop on Integrating Artificial Intelligence and Assistive Technology* (pp. 40–47). Madison, WI: AAAI Press.

Platt, L., Andrews, G., & Howie, F. M. (1980). Dysarthria of adult cerebral palsy: II. Phonemic analysis of articulation errors. *Journal of Speech & Hearing Research, 23,* 41–55.

Platt, L., Andrews, G., Young, M., & Quinn, P. (1980). Dysarthria of adult cerebral palsy: I. Intelligibility and articulatory impairment. *Journal of Speech & Hearing Research, 23,* 28–40.

Rosenbek, J., & LaPointe, L. (1978). The dysarthrias: Description, diagnosis, and treatment. In D. Johns (Ed.), *Clinical management of neurogenic communicative disorders* (pp. 251–310). Boston: Little-Brown and Company.

Rounsfell, S., Zucker, S. H., & Roberts, T. G. (1993). Effects of listener training on intelligibility of augmentative and alternative speech in the secondary classroom. *Education and Training in Mental Retardation, 26,* 296–308.

Schlenck, K., Bettrich, R., & Willmes, K. (1993). Aspects of disturbed prosody in dysarthria. *Clinical Linguistics and Phonetics, 7,* 119–128.

Shein, F., Brownlow, N., Treviranus, J., & Parnes, P. (1990). Climbing out of the rut: The future of interface technology. In B. A. Mineo (Ed.), *ASEL Applied Science and Engineering Laboratories: Augmentative and Alternative Communication in the Next Decade* (pp. 17-19). Wilmington, DE: University of Delaware Press.

Smith, M. M. (1994). Speech by any other name: The role of communication aids in interaction. *European Journal of Disorders of Communication, 29,* 225–240.

Traunmuller, H. (1981). Perceptual dimension of openness in vowels. *Journal of the Acoustical Society of America, 69,* 1465–1475.

Treviranus, J., Shein, F., Haataja, S., Parnes, P., & Milner, M. (1991). Speech recognition to enhance computer access for children and young adults who are functionally non-speaking. In *Proceedings of RESNA 14th Annual Conference* (pp. 308–310). Washington, DC: RESNA Press.

Vance, J. E. (1994). Prosodic deviation in dysarthria: A case study. *European Journal of Disorders of Communication, 29,* 61–76.

Vanderheiden, P. J. (1985). Writing aids. In J. G. Webster, A. M. Cook, W. J. Tompkins, & G. C. Vanderheiden (Eds.), *Electronic aids for rehabilitation* (pp. 262–282). London: Chapman and Hall.

Wit, J., Maassen, B., Gabreels, F. J. M., & Thoonen, G. (1993). Maximum performance tests in children with developmental spastic dysarthria. *Journal of Speech and Hearing Research, 36,* 452–459.

Yorkston, K. M., & Beukelman, D. R. (1981). *The Assessment of Intelligibility of Dysarthric Speakers.* Austin, TX: PRO-ED.

Yorkston, K. M., Beukelman, D. R., Strand, E. A., & Bell, K. R. (1999). *Management of motor speech disorders in children and adults.* Austin, TX: PRO-ED

Yorkston, K. M., Hammen, V. L., & Dowden, P. (1991). Index of contextual intelligibility: A perceptual analysis of intelligible versus unintelligible productions in severe dysarthria. In C. A.

CARTER & MAXWELL

Carter, M., & Maxwell, K. (1998). Promoting interaction with children using augmentative communication through a peer-directed intervention. *Journal of Disability, Development and Education, 45,* 75–96. One factor that is critical to the successful integration of children using augmentative and alternative communication (AAC) systems is establishing interaction with peers. AAC systems have the potential to increase the opportunities for interaction, but successful social interaction is dependent on a range of factors including the communicative knowledge, skills, and attitudes of partners. The present study attempted to increase the quantity of social interaction in classroom settings between children (aged 5 to 9 years) using AAC systems and their peers. A multifaceted intervention was directed at communicative partners, and most particularly, peers. The study was 15 weeks in duration and utilised a multiple baseline across subjects design. The study demonstrated the effectiveness of an intervention to increase social interaction.

Reprinted with permission of Carfax Publishing, Oxfordshire, UK.