The Impact of Cerebral Palsy on the Intelligibility of Pitch-based Linguistic Contrasts

Valter Ciocca, Tara L. Whitehill and Ma Ka Yin, Joan

Division of Speech & Hearing Sciences, University of Hong Kong, Hong Kong

Abstract This study investigated the perception of three Cantonese level tones produced by speakers with dysarthria associated with cerebral palsy. Four speakers with dysarthria were selected on the basis of their distinctive patterns of fundamental frequency (F0) values observed in the level tones they produced, which showed errors in either F0 level or, F0 contour, or both. Monosyllabic words which contrasted in tone level were used as stimuli in an identification task. Five expert listeners identified the tones among the six Cantonese contrastive tones. Results showed that the tones produced by the dysarthric speakers were predominantly perceived as level tones; although a majority was perceived as the wrong level tone. The most important finding is that only the level tones produced by dysarthric speakers could be identified as rising or falling contour tones. The frequent perceptual confusion among the level tones, and the perception of contour tones, shows that a disorder in the production of pitch-based linguistic contrasts can have a substantial impact on the communication abilities of individuals with cerebral palsy.

Introduction

Cerebral palsy is a neurological disorder of posture and movements that usually result from the damage of the brain during intrauterine development (Pellegrino, 2002). The intelligibility of individuals with cerebral palsy is often compromised because of dysarthria, the most common speech disorder found in individuals with cerebral palsy (Workinger and Kent, 1991). All components of the speech production system, including respiration, phonation, articulation, resonance and prosody, can be affected by motor disorders associated with dysarthria (Byrne, 1959; Darley et al., 1969a; Duffy, 1995; Platt et al., 1980).

Although most studies of dysarthric speech have been carried out on English-speaking populations, there have been a number of investigations on the effects of dysarthria on speech production in other languages—such as Japanese (Kumai et al., 1978; Fukusako et al., 1983), German (Ziegler et al., 1988), French (Gentil, 1992), and Cantonese (Whitehill and Ciocca, 2000ab). These studies are important because they provide an evidence for the disorder which are common to all languages, as well as language-specific effects related to the characteristics of specific phonological systems.

One of the central differences between Cantonese and English phonology is that Cantonese is a tonal language, in which lexical tones affect the meaning of the monosyllabic units with which they co-occur. In the Cantonese spoken in Hong Kong, there are six contrastive tones (Fok Chan, 1974; Bauer and Benedict, 1997): high-level (55), high-rising (25), mid-level (33), low-falling (21), low-rising (23) and low-level (22). These tones are mainly contrasted on the basis of their fundamental frequency (F0) level and contour, which are perceived by speakers in terms of corresponding pitch patterns (Fok Chan, 1974; Vance, 1976). F0 values associates with each lexical tone are not absolute, but depend on the F0 range of individual speakers (Fok Chan, 1974; Bauer and Benedict, 1997). For example, female speakers have a higher average F0 than male speakers; consequently a high level tone produced by a male speaker could have similar F0 values to a low level tone produced by a female speaker.

During the normal speech production, the speakers of tonal languages alter the tension of the vocal folds and the volume of air flowing from the lungs through the glottis in order to produce different tones (Ladefoged, 1993). By contrasts, dysarthric speakers may not be able to control the respiratory and the laryngeal mechanisms for the correct production of pitch-related linguistic information. Darley et al. (1969b) reported that phonatory stenosis and incompetence, which were due to abnormalities in the laryngeal mechanism, affected different types of dysarthria. Jacques et al. (1985) observed a reduction in laryngeal control in cerebral palsy patients with congenital spasticity. Theodoros and Murdoch (1994) reported
a high incidence, and a wide range, of laryngeal dysfunction in dysarthria associated with closed-head injury. Laryngeal and respiratory dysfunction is likely to cause disorders in the production of pitch-related features of language production. Indeed, Bunton et al. (2000) conducted a study of the prosodic characteristics of English dysarthric speech and found that the F0 range was restricted relative to that of the normal speech (see also Le Dorze et al., 1994). Recent studies have also reported disorders in the production of pitch-based lexical tone contrasts in Cantonese. A narrower range of variation in F0 has been reported for Cantonese speakers with dysarthria caused by Parkinson’s disease (Wong, 1999). He found that dysarthric patients spoke with a more restricted pitch range, and that their lexical tones were identified less accurately than tones produced by non-dysarthric speakers. Whitehill and Ciocca (2000ab) investigated the intelligibility of words produced by Cantonese dysarthric speakers with cerebral palsy. They found that the tone-level contrast (involving errors among level tones) was the 2nd most problematic phonetic contrast associated with speech intelligibility; the tone-contour contrast, reflecting the confusion among contour tones, was the 7th most problematic contrast. The tone level contrast was vulnerable across speakers with various severity levels, and accounted for a relatively large proportion of the variance particularly for speakers with a mild intelligibility deficit.

In addition to perceptual studies, acoustic analysis studies have provided additional evidences of disorders in tone production by Cantonese dysarthric speakers. Whitehill et al. (2000) studied the fundamental frequency (F0) patterns of level tones produced by Cantonese dysarthric and non-dysarthric speakers. They found that both male and female dysarthric speakers showed overall higher F0 values and greater variability in their tone production. Five abnormal patterns were identified as follows: (i) excessive variability, meaning overlaps in standard error bars between tones, (ii) excessively falling frequencies, (iii) lowering of the high level tone, (iv) rising tone contour for level tones, and (v) abnormal contour patterns.

The purpose of the present study was to explore the relationship between the F0 characteristics and the intelligibility of lexical tone production by Cantonese dysarthric speakers. The following hypotheses were made for the present study: i) abnormal F0 patterns in lexical tones produced by dysarthric speakers should result in a lower percentage of correct identification; ii) the error patterns in the identification of tones should be consistent with the types of abnormal F0 patterns (for example, a level tone produced with a falling F0 should be perceived as falling tone).

**Methods**

**Subjects**

Four dysarthric speakers, who participated in Whitehill and Ciocca’s (2000ab) study, were selected. The age of the four dysarthric subjects ranged from 18 to 36 (mean age 25.7). The type of cerebral palsy included spastic (1), athetoid (2) and mixed (1). Their intelligibility scores ranged from 28 to 92% (for a detailed report on the calculation of the intelligibility scores, see Whitehill and Ciocca, 2000b). One non-dysarthric male was included as a control speaker. All the participants were native speakers of Cantonese; they had normal hearing (≤20 dBHL at octave frequencies between 250 and 8000 Hz), normal oral-peripheral structure, and normal language ability (screened by the Cantonese Aphasia Battery; Yiu, 1992).

Five undergraduate students on the fourth year in Speech and Hearing Sciences at the University of Hong Kong served as listeners. They were all native Cantonese speakers; they all had training in tone perception and had previous exposure to dysarthric speech. Their tone perception abilities were screened using twenty-four single monosyllabic words (four for each of the six tones) produced by two non-dysarthric speakers (one male and one female). All the listeners were able to correctly identify at least two out of four stimuli per tone, and on average, they achieved 92% accuracy in the screening session.

**Stimuli and Procedures**

The following twelve monosyllabic words with four tokens for each of the three level tones were chosen as the stimuli: /kwaai55/; /kwaai33/; /ham55/; /ham33/; /soeng55/; /soen22/; /se55/; /tse22/; /fong33/; /fan22/; /tsiu33/; /sin22/. These were produced as part of a list of 75 words included in the Cantonese Single-Word Intelligibility Test (CSIT; Whitehill and Ciocca, 2000b). Single-word production was chosen since some dysarthric speakers had difficulties in producing longer utterances. Speech samples were collected in a sound-attenuated room, with a Sony TCD-D3 DAT recorder and a Bruel & Kjaer (4003) low-noise unidirectional microphone. A 10 cm mouth-to-microphone distance was maintained during the recording. The words in the CSIT were produced in randomized order within the word list. Repetition after the experimenter was used as the elicitation method for the collection of the speech samples.

After recording, the utterances were stored onto an Apple PowerMacintosh 7100 computer as separate files, using a DigiDesign Audiomedia II DSP card with low-pass filtering at 22 kHz and a sampling rate of 44.1 kHz. The loudness of each sound file was perceptually adjusted by the third author as follows by first rating the loudness of each stimulus on a 5-point scale (‘1’=‘soft’; ‘5’=‘loud’). Stimuli rated as soft (1 or 2 on the scale) were normalized using the Sound Designer II software, and then the loudness of the stimuli rated as loud (4 or 5 on the scale) was attenuated so that their loudness matched the loudness of the normalized stimuli.

The dysarthric speakers were selected among the speakers in Whitehill and Ciocca (2000b) and Whitehill et al. (2000), according to the distinctive mean F0 patterns associated with their production of the high level, mid level and low level tones. For speaker 1, the F0 pattern of the three level tones
resembled the normal pattern (`Normal F0` type). Speaker 2 produced an overlap in the F0 range for level tones: the F0 values for the high level and low level tones overlapped, and those of tone 33 were the highest (`Overlapping F0` type). The third speaker produced an excessively falling F0 contour for all level tones (`Falling F0` type). Speaker 4 produced both excessively falling F0 patterns and an overlap between the mid level and the low level tones (`Mixed F0` type).

The experiment was carried out in a sound-attenuated room (IAC sound-proof booth), with the speech stimuli presented to the listeners through Sennheiser HD545 headphones, connected to an Apple PowerMacintosh 7100 computer. A HyperCard program was used for running the experiment. Each test item was presented twice within each trial. Each of the twelve stimuli was repeated five times, resulting in a total of 300 stimuli (20 stimuli for each tone for each of five speakers). The trials were divided into two listening blocks, one for the control and another for the dysarthric speakers.

Stimuli within blocks were randomized automatically by the HyperCard program. The listener had to identify the lexical tone of each word by clicking one of the six buttons named Tone 55, Tone 25, Tone 33, Tone 21, Tone 23, and Tone 22 on the computer screen. Listeners were informed that they should focus on the identification of lexical tones when listening to the dysarthric productions, rather than identifying the intended words in Cantonese, since the tones and the segments produced did not necessarily sound like the normal production for a given word. For the same reason, all listeners listened to the control utterances first, in order to get familiar with the task. The two blocks took about one hour to complete.

### Results

The confusion matrices for the control and the dysarthric speakers are shown in Table 1. For the control speaker, tone 55 (high level) was accurately identified, but the mid level tone (tone 33) was often perceived as tone 55, probably because the mean F0 end point for this speaker was relatively close to that of tone 55. Interestingly, and unexpectedly, for this speaker tone 22 was consistently misidentified as tone 33.

For the dysarthric speakers, tones 55 and 33 were more accurately identified than tone 22, and were frequently confused with each other. Tone 22 was still the most difficult to perceive correctly. Tone 22 was confused with tone 33, but also with tone 55 about 18% of the times (more than twice as often as for the tone 22 produced by the control speaker). The level tones produced by the control speaker were not perceived as contour tones. By contrast, the dysarthric level tones were sometimes identified as contour tones. For instance, tone 33 was identified as the high rising tone (tone 25) about 6% of the times, and as a contour tone about 9% of the times. Tone 22 was identified as a contour tone about 6% of the times, and was misperceived as either tone 23 or tone 21. Tone 55 was also perceived as contour, but less frequently than the other level tones.

Table 2 illustrates the response patterns for the tones produced by each dysarthric speaker. As expected the pattern of responses for the Normal F0 speaker was similar to that of the control speaker. Listeners perceived his high level, mid level and low level tones with 88, 58 and 4% accuracy, respectively. Only two out of 300 stimuli were misidentified as a contour tone (tone 23); all other errors are in relation to the confusion among level tones.

For stimuli produced by the speaker with Overlapping F0, there were frequent confusions among the three level tones, as expected from the mean F0 pattern produced by this speaker. The fact that tone 33 was the most frequently identified as tone
55 is also consistent with the fact that tone 33 had the highest mean F0 for this speaker. A number of utterances were identified as contour tones, mainly as the low rising tone (23). For the speaker with Falling F0 productions, tone 33 was the most accurately identified (52% correct), followed by tone 22 (36%) and tone 55 (26%). Most noticeably, tone 33 was identified as the high rising tone 25% of the times. Other contour tone responses involved tone 23 (8 responses) and tone 21 (5 responses). The tones produced by the speaker with Mixed F0 patterns were accurately identified as high level and mid level tones 65% of the times. Tone 22 was identified less accurately (20% correct). Relatively more instances of low falling tone responses (15 in total) were given for intended tones 33 and 22 produced by this speaker than for tones produced by the other dysarthric speakers.

Discussion

As expected, tones produced by dysarthric speakers were more poorly identified than tones produced by the control speaker overall; confusions occurred among all level tones, as expected due to the overlap in F0 values among the three level tones produced by these speakers. There were many more tone level errors than tone contour errors for both groups. This finding implies that, although they had difficulties in producing tones with appropriate F0 level, the dysarthric speakers were generally able to preserve F0 information associated with the level tones in speech production. Nonetheless, the most interesting aspect of the data is the finding that contour tone responses occurred relatively frequently for tones produced by dysarthric speakers. The percentage of tone contour responses for the dysarthric speakers ranged from about 3% (tone 55) to about 9% (tone 33). This finding can be explained by the fact that many of the level tones produced by dysarthric speakers had in fact relatively large F0 changes, unlike the level tones produced by the control speaker. By contrast, there was only one tone contour response (tone 23) out of 300 judgments for the control speaker.

Among the four dysarthric speakers, the Normal F0 speaker had the highest correct identification rate for tones 55 and 33 among dysarthric speakers, as expected. Moreover, the relatively frequent identification of tones 33 and 22 as tone 55 is likely to be due to his higher than normal overall F0, which is typical of dysarthric speakers (Whitehill et al., 2000). For the Falling F0 speaker, this F0 pattern did not produce a large number of tone 21 responses as might have been expected. Instead, most responses consisted of the selection of a level tone. One of the possible explanations is that, since there is only one falling tone in Cantonese, some abnormally falling F0 contours may not be perceived as a falling tone if they are not occurring within the expected F0 range, or if their F0 drop exceeds what is expected for a typical falling tone. Tones produced by the Overlapping F0 speaker were the most poorly identified, likely because of the disruption in the relative level of the three level tones. Tone 33, which was produced with F0 values above those of tone 55, was most often perceived as tone 55 (52.19%). Tones 55 and 22, which overlapped in F0 values, had similar identification patterns. These findings are consistent with the idea that the perception of Cantonese level tones depends mainly on their relative F0 levels.

The identification errors for the controlled speaker showed that tone 55 was the most robust among the three level tones they produced. The confusion between tones 33 and 22 was expected since their F0 values for this speaker were very similar, which is typical for the normal Cantonese speakers (Fok Chan, 1974; Whitehill et al., 2000). Previous studies have also reported that tones 33 and 22 were perceptually confused when heard as single words in an identification task (Fok Chan, 1974; Wong and Diehl, 2003). Listeners consistently identified tone 22 as tone 33 at most of the time due to the unclear reasons. Since the production of level tones by this speaker was likely to be atypical, an obvious extension of this study would compare the performance of the dysarthric group with that of the controlled group including several non-dysarthric speakers.

This study showed that the perception of tones produced by dysarthric speakers can be problematic, and that a disruption in relative F0 level can have substantial impact on communication because it can adversely affect the intelligibility of tones. For non-dysarthric speakers, inaccurate productions in the relative F0 level of level tones can be compensated in conversational speech by the tonal context through tone normalization (Leather, 1983; Wong and Diehl, 2003, Francis et al. 2003). However, due to the excessive variability in the production of pitch-related linguistic contrasts and the shortness of phrases commonly observed for dysarthric speakers (Darley et al., 1969a; Bunton et al., 2000), a disruption in relative F0 level may still have significant impact on their communication abilities.

Acknowledgement We would like to thank: Alice Lee, for her advice and support in the planning and the data collection for this study; Raymond Wu and Donald Chan, for their technical support in the listening procedures; all the speakers and the listeners who participated in this study; Yoshitaka Nakajima for making helpful comments on the first draft of this manuscript.

References


Fok Chan AYY (1974) A Perceptual Study of Tones in Cantonese. University of Hong Kong Press, Hong Kong


Received: September 6, 2004
Accepted: September 30, 2004
Correspondence to: Valter Ciocca, PhD, Division of Speech & Hearing Sciences, University of Hong Kong, 5/F, Prince Philip Dental Hospital, 34 Hospital Road, Hong Kong, Hong Kong SAR, China
Phone: +852–2859–0581
Fax: +852–2559–0066
e-mail: vciocca@hkusua.hku.hk