For inclusion in N. Miller & A. Lowit (Eds.), *Motor Speech Disorders: A Cross Language and Bilingual Perspective (or Communication disorders across languages: Acquired motor speech disorders)*. Chapter 9B

Motor speech disorders in Chinese

Tara L. Whitehill and Joan K-Y. Ma

Corresponding author: Tara L. Whitehill. Corresponding address: Division of Speech & Hearing Sciences, University of Hong Kong, Pokfulam Road, HONG KONG
MOTOR SPEECH DISORDERS IN CHINESE

Chinese, spoken by approximately 20% of the world’s population, is the most commonly spoken language in the world (Fung, 1990). There is some debate regarding whether varieties of spoken Chinese should be considered dialects or separate language. Here, we regard them as separate languages. The various languages of Chinese are considered united by a common written system, although this characterization is problematic for Cantonese, which has many colloquial expressions with no written form (Bauer & Benedict, 1997). In this chapter, we focus on two of the most common Chinese languages, Cantonese and Mandarin (Putonghua). Cantonese is spoken in Hong Kong, Southern China and many overseas Chinese communities. Mandarin, also known as Standard Chinese, is the official national language of China and Taiwan, and is one of the four official languages of Singapore.

Cantonese

Detailed descriptions of Cantonese phonology can be found in Bauer and Benedict (1997), Cheung (1986) and Zee (1991). There have been active debates about several aspects of Cantonese phonology, including (a) the treatment of final glides/diphthongs, (b) the nature of the consonants /kw/ and /kwʰ/, and (c) the number of lexical tones. In this chapter, we have generally adopted the positions of Bauer and Benedict (1997).

Syllable structure. Traditionally, Chinese syllables have been described using an onset, a rime (comprising an obligatory vowel and an optional coda) and a tone. We adopt here an alternate model of the syllable structure (Bauer & Benedict, 1997), which takes account of the two permissible syllabic nasals, /m/ and /ŋ/. The structure is: (C1) V1 or Cn (C2 or V2), where C1 = initial consonant, V1 = vowel, Cn = syllabic nasal consonant, C2 = final consonant, and V2 = ending vowel (second portion of diphthong); portions in parentheses are optional. Over 98% of Cantonese syllables are CV or CVC in structure (Wang, 1941, cited in Lau & So, 1988). Tone, which is carried on the vowel portion of the syllable, is obligatory.
Phonology. Cantonese has 19 initial consonants: [p, pʰ, t, tʰ, k, kʰ, ts, tsʰ, f, s, h, m, n, ɲ, l, w, j, kw, kwʰ]. The initial nasal /ɲ/ is optionally deleted in contemporary Cantonese, and there is a free variation between initial /n/ and /l/. There is some debate about the segments /kw/ and /kwʰ/, which are considered either as coarticulated unitary phonemes or as clusters. The six final consonants are [p, t, k, m, n, ɲ]. The final plosives are unreleased. There are eight primary vowels, [i, y, u, ɛ, œ, ɔ, a, ɐ], each of which has allophonic variations. The short vowel, /ɐ/, only appears with a final consonant, whereas the other vowels can form a rime component independently. In addition, Cantonese has ten diphthongs: [ai, ui, ei, ɐi, ɔi, ɛi, au, ɔu, ɐu, iu, œy]. Cantonese is a lexical tonal language, where variations in fundamental frequency are used to differentiate minimal word pairs that are not distinctive by segmental information. There are six contrastive tones in contemporary Cantonese: high level (55), high rising (25), mid level (33), low falling (21), low rising (23) and low level (22). There are several methods to describe Cantonese tones; here, we have combined traditional verbal descriptors with the numeric system developed by Chao (1947), with a modification to the tone value of the high-rising tone from 35 to 25, which has been found to be a better description of the fundamental frequency contour (Ma, Ciocca & Whitehill, 2006).

Stress and rhythm. In contrast to English, which is stress-timed, Cantonese is syllable-timed. Stress is not phonemic in Cantonese.

Syntax/morphology considerations. A thorough description of Cantonese syntax and morphology can be found in Mathews and Yip (1994) and a shorter summary in Fung (2009). To our knowledge, possible interactions between syntax/morphology and phonology have not received any attention in the literature on Chinese motor speech disorders.

Features of particular interest. As most of the literature about motor speech disorders has focused on English speakers, here we note the features of Cantonese which
are most in contrast with English. These are the features most likely to be of interest in investigating possible language-specific aspects of motor speech (as well as other speech) disorders. First is the tonal nature of Cantonese. Second is the relatively simple syllable structure (i.e., with /kw/ and /kwʰ/ being the only clusters or, in some cases, considered as unitary phonemes, and 98% of syllables either CV or CVC). Third, the relatively small fricative system (only three fricatives, only one of which is produced intra-ornally). Fourth, the fact that Cantonese has an aspirated versus unaspirated contrast for plosives, as opposed to the voiced-voiceless contrast in English (both contrasts involve differences in voice onset time; see Clumeck, Barton, Macken & Huntington, 1981). Fifth, that Cantonese is not stress-timed. Finally, the logographic versus alphabetic nature of the written language (which raises issues for treatment, in particular).

Mandarin

Detailed descriptions of Mandarin phonology can be found in Ladefoged and Maddieson (1996), Hashimoto (1970), Lee and Zee (2003) and Hua (2002). Controversies regarding Mandarin phonology have included (a) the characterization of /ɻ/ (sometimes described as /ɾ/) and /χ/ (sometimes described as /χ/); (b) the surface value of certain vowels, given their allophonic variations; and (c) whether the vowels /i/ and /u/ should be considered as semivowels when occurring in diphthongs or triphthongs (Hua 2002, p. 42).

Syllable structure. The syllable structure of Mandarin is (C) V (N). That is, initial consonant is optional, vowel is obligatory, and final consonant, which can only be a nasal consonant, is optional. As with Cantonese, tone is obligatory and is carried on the vowel portion of the syllable.

Phonology. Mandarin has 22 initial consonants: [p, pʰ, t, tʰ k, kʰ m, n, ŋ, ts, tsʰ, tʃ, tʃʰ, tɕ, tɕʰ, f, s, ş, ç, x, i, l]. There are only two final consonants, [n, ŋ]. There are nine simple vowels, [i, y, e, æ, ə, a, ɤ, u, o], nine diphthongs, [ae, ao, ei, ou, ia, ie, ua, uo, ye], and four
triphthongs, [iɑo, iou, uae, uei]. Mandarin has four tones, described as high level (55), rising (35), falling-rising (214), and high falling (51); the numbers again refer to the numeric system of Chao (1930).

**Stress and rhythm.** Unlike Cantonese (but as in English), Mandarin is a stress-timed language. Weak stress (also termed weak syllable or neutral tone) is a particular characteristic of Mandarin, and has attracted considerable debate, in terms of the relationship between weak syllable and tone, and relevant phonological and morphological rules (Hua, 2002).

**Syntax/morphology considerations.** A detailed description of Mandarin grammar can be found in Chao (1968); a more succinct and recent description is provided by Fung (2009). As with Cantonese, we are not aware of any reports of interactions between syntax/morphology and phonology in the literature on motor speech disorders in Mandarin speakers.

**Features of particular interest.** Some of the same features mentioned above for Cantonese are also of interest when contrasting Mandarin with English. Namely, the tonal nature of Mandarin, the relatively simple syllable structure, the aspiration contrast for plosives, and the logographic nature of the written language. Several additional features are of interest when contrasting Mandarin phonology with both English and Cantonese. First is the relatively restricted final consonant system and the fact that the only permissible finals are two nasals. Second, the fricative/affricate system is relatively rich in comparison to Cantonese, with retroflex fricative/affricate, alveolo-palatal fricative/affricate and velar fricative. Third, the place of articulation involves both retroflex and alveolar-palatal, which does not exist in Cantonese. Again, these are features that should be of particular interest when investigating possible language-specific influences on motor speech disorders.

**Studies of motor speech disorder in Chinese**
Whitehill (2010) provided a recent review of studies of motor speech disorder in Chinese, which focused on adults with Parkinson’s disease (PD), and teenagers or young adults with cerebral palsy. Here, we focus primarily on acquired disorders. This chapter also takes a more applied focus, reviewing available clinical materials as well as treatment studies.

**Characterization of motor speech disorders.** There are limited studies focusing on general characterization of motor speech disorders in Chinese. Using the classic Mayo clinic approach (Darley, Aronson and Brown, 1969a, 1969b, 1975), Whitehill Ma and Lee (2003) investigated the perceptual characteristics of hypokinetic dysarthria in Cantonese speakers. The findings were largely similar to those found for English (e.g., Darley et al., 1975) and Japanese (Fukusake, Monoi, Tatsume et al., 1983). Interestingly, lexical tone production was found to be relatively robust in this group of speakers, particularly in comparison with the dimension of ‘monotone’. To our knowledge, no similar studies have been done for other types of dysarthria in Cantonese or Mandarin.

Kwan (1998) investigated acoustic variables of speech in one Cantonese speaker with apraxia and three speakers with ataxic dysarthria. The acoustic measures included variables associated with vowels and diphthongs, temporal organization, pause duration, and the intensity of syllable-initial consonants and vowels. The findings were generally consistent with those of previous studies of English speakers with similar motor speech disorders. The fundamental frequency (F0) contours of tones were deviant for all four speakers, compared with those of non-impaired speakers.

**Assessment and treatment materials.** There are few published tests or materials for the assessment of motor speech disorders in Chinese speakers. The National Institutes of Health Stroke Scale (NIHSS) is a well-known 15-item scale to evaluate patients following stroke (Brott, Adams Jr, Olinger et al., 1989). The scale contains one item on facial palsy (rated on a 4-point scale), and one item on dysarthria, which involves the patient reading or repeating a list of words, and is rated using a 3-point scale. The NIHSS has been translated
and validated for Cantonese (Cheung, Lyden, Tsoi et al., 2010) and Mandarin (Sun, Chiu, Yeh and Chang, 2006; Cheung, et al., 2010).

Whitehill and colleagues developed a series of materials to be used in evaluating Cantonese speakers with dysarthria. These materials include a single-word list for evaluating phonology (Whitehill, 1994); a contrastive single-word intelligibility test (Whitehill, 1995, Whitehill & Ciocca, 2000a) that is based on the test developed by Kent, Weismer, Kent and Rosenbek (1989); and a Cantonese sentence intelligibility battery (Whitehill, 2003; Whitehill, Ciocca and Yiu, 2004), based on the sentence portion of the AIDS (Yorkston & Beukelman, 1981). These materials have been primarily used in research studies; they are available upon request from the first author but have not been published. Most clinicians working in Hong Kong evaluate dysarthria based on principles and procedures outlined by Duffy (2005). Apaxia is generally evaluated using an informal translation of the Apraxia Battery for Adults (Dabul, 2000). The validity of the non-standardized translation has not been systematically determined. However, the Battery seems generally appropriate for Cantonese, for tasks such as increasing word length. Yiu (1992) developed a Chinese (Cantonese) version of the Western Aphasia Battery (WAB; Kertesz, 1982), which includes a 15-item section for the evaluation of apraxia.

A single-word intelligibility test was developed in Taiwan for Mandarin speakers with dysarthria (Liu, Tseng & Tsao, 2000). The test was also based on that developed by Kent and colleagues (Kent et al., 1989). Liu and colleagues (2000) describe the use of the test with young men with cerebral palsy but the test is presumably suitable for use with other groups of Mandarin speakers with dysarthria. We are not aware if the test is being used clinically, but it appears to be primarily a research tool.

There are few published treatment materials available for Chinese speakers with motor speech disorders. The website for Lee Silverman Voice Treatment (LSVT®) lists the availability of a Cantonese version of the DVD for treatment materials; this is the only material we are aware of.
Treatment studies. Reports of treatments studies with Chinese speakers with dysarthria are, unfortunately, sparse. A recent study by Whitehill and colleagues described the use of Lee Silverman Voice Treatment (LSVT®) with a group of Cantonese patients with PD (Whitehill, Kwan, Lee & Chow, accepted). The primary focus of the study was to investigate the effect of LSVT® on lexical tone errors. The speakers made significant improvement on a number of measures already well-documented to improve following LSVT® treatment in English speakers, such as loudness (Ramig, Sapir, Fox & Countryman, 2001) and vowel articulation (Sapir, Spielman, Ramig, Story & Fox, 2007). However, there was no significant improvement in lexical tone. This confirmed the findings of an earlier pilot study with a smaller group of Cantonese speakers with PD, using treatment based on the principles of LSVT® (Whitehill & Wong, 2007). The authors discussed the relatively intact nature of tone pre-treatment as well as a possible dissociation between tone and intonation. A more detailed discussion on the interaction between intonation and tone in speakers with dysarthria is included in a later section (Language specific studies).

The effects of surgical treatment (bilateral subthalamic nucleus deep brain stimulation, STN-DBS) on speech in a group of Mandarin-speaking patients with PD were investigated by a team of researchers in Beijing (Xie, Zhang, Zheng et al. 2011). The results confirmed those of several previous studies of deep brain stimulation surgery for this population. That is, while the procedure improved motor abilities, there was effectively no improvement in speech performance. The study did not focus on language-specific aspects of performance, pre- or post-surgery.

Language Specific Studies

Tone. One of the main characteristics of Chinese languages (Cantonese and Mandarin) is that they are tonal languages. In tone languages, F0 variation at the syllabic level is used to mark semantic meaning, while F0 changes at the sentential level are used to mark intonation. The production of lexical tone in Cantonese speakers with dysarthria has been the focus of investigation in two etiological groups, Parkinson’s disease and cerebral palsy, for different reasons. Speakers with PD are characterized by monotonous speech and
reduced pitch variation within sentences due to rigidity, while speakers with cerebral palsy
tend to display extreme variability related to reduced control of the laryngeal mechanism
(Darley, Aronson & Brown, 1969b; Jacques, Rastatter & Sullivan, 1985). As a result of these
different characteristics, both groups are susceptible to errors in tone production.

Wong and Diehl (1999) investigated the production and perception of Cantonese
tones produced by a PD speaker and a non-impaired speaker. They reported that the PD
speaker had a more restricted tonal space, as defined by the pitch range of all tones, when
compared with the health control speaker. Additionally, the tones produced by the PD
speaker were less accurately identified than those of the non-PD speaker. However, they
provided little detail about their methodology. Also, their study used only one PD speaker.
As speech produced by individuals with dysarthria is known to be highly heterogeneous
(Lowit-Leuschel & Docherty, 2001), it is questionable whether the pattern reported in Wong
and Diehl (1999) was representative of PD speakers. Using perceptual analysis, Whitehill,
Ma and Lee (2003) established a perceptual profile of the speech characteristics of
Cantonese PD speakers. The dimension of “tone distortion” was included specifically for the
linguistic property of Cantonese. They found that tone production was relatively unaffected
in Cantonese PD speakers. Ma (2009) investigated the acoustic pattern of the Cantonese
lexical tones produced by five speakers with PD, with the target tones embedded at three
different positions (initial, medial and final) of a five-syllable question or statement. The
results showed that speakers with PD contrasted the six lexical tones in a similar manner
compared with control speakers across positions and intonations, except at the final position
of questions. Significantly lower fundamental frequency (F0) values were found towards the
end of the syllable at the final position of questions for the speakers with PD than for the
control speakers. This showed a different pattern of interaction between intonation and tone
for the speakers with PD and the control speakers. In investigating the effect of medication
on speech in Mandarin speakers with PD, Tseng (2000) included tonal contrast as one of the
measures. Using perceptual analysis, the results showed that the lexical tones produced by
ten PD speakers were very similar to those of the control speakers.
A series of studies were conducted to investigate lexical tone production by Cantonese speakers with cerebral palsy and perception of tone production by non-dysarthric speakers (Ciocca, Whitehill & Chow, 2000; Ciocca, Whitehill & Ma, 2004; Ciocca, Whitehill & Ng, 2002). The F0 patterns of the monosyllabic tone targets were analysed acoustically and the results showed excessive variability in the F0 patterns by speakers with cerebral palsy. The six tones showed a much larger degree of overlap than in non-dysarthric speakers.

Abnormal patterns in both tone level and tone contour were observed. The F0 level of the high-level tone was found to be lower than for non-dysarthric speakers. Additionally, falling F0 patterns were found for both level tones and rising tones, while rising F0 patterns were observed in some falling tone production. These errors patterns had a significant impact on perception, with level tones being perceived as rising or falling, and errors perceived in the perception of tone level. Jeng, Weismer & Kent (2006) conducted a similar investigation of tone production and its perception in Mandarin speakers with cerebral palsy. They found that the high-rising and low falling rising tones were produced less accurately than the high level and the high falling tone but commented that, in general, the F0 contours of the four tones were to a large extent retained by the speakers. Instead, the acoustic analysis showed that the most significant problem in lexical tone production in this group of speakers was related to the precision of the F0 contour.

**Voice Onset Time.** Another specific feature of Chinese languages is that, instead of contrastive voicing, plosives and affricates are contrasted by aspiration. This applies to both Cantonese and Mandarin. Voice onset time (VOT) serves as an acoustic correlate and perceptual cue of the aspiration contrast, as for the voicing contrast in English. One study has examined the VOT of plosives in Mandarin speakers with dysarthria (Tseng, 2000). Ten speakers with PD were asked to produce a series of words contrasting in both tone and aspiration. The results of the acoustic analysis showed speakers with PD showed a larger number of VOT overlaps than control speakers, reducing the contrast between aspirated and unaspirated consonants. Tseng (2000) hypothesized that the VOT overlap was related to reduced coordination of the speech mechanism in speakers with Parkinson’s disease.
Interestingly, some speakers in the same study showing no VOT overlap before L-dopa medication were found to have VOT overlap after L-dopa. Another study investigating the effect of various acoustic parameters on intelligibility in 20 Mandarin speakers with cerebral palsy showed longer VOT for unaspirated stops than in control speakers (Liu et al., 2000).

**Consonant Cluster.** Reduction of consonant clusters /kw/ and /kw^h/ was found to be one of the most common errors in the manner of articulation among Cantonese individuals with cerebral palsy (Whitehill & Ciocca, 2000b). The consonants /kw/ and /kw^h/ were realized as either [k] or [k^h], which are considered acceptable variations in Cantonese (in some contexts), or as [w], which is regarded as an error. The simplification of the manner of articulation in these consonant clusters reflected the nature of the dysarthrias associated with cerebral palsy, whereby individuals lack skills in precise articulatory movement rather than an inability to contrast phonemic prominence (Platt, Andrews & Howie, 1980). The findings also have implications for the controversy surrounding the nature of the /kw/ and /kw^h/ phonemes in Cantonese (consonant cluster versus unitary phoneme), as the error patterns showed cluster reduction (i.e., /kw/ → [k]).

**Theoretical implications**

Speech prosody has been found to be one of the main impairments in Chinese speakers with dysarthria (e.g. Ciocca, Whitehill & Lam, 2001; Whitehill, Ma & Lee, 2003), as in other languages. As F0 is used to mark both intonation (at sentential level) and lexical tone (at syllabic level), it was traditionally believed that an impairment in intonation marking in speakers with dysarthria would unavoidably be generalized to the production of lexical tones. This depends on the whether there are separate control for tone and intonation, and the degree of interaction between tone and intonation. There is little evidence of separate physiological control for tone and intonation. Vance (1976) hypothesized that there might be separate control mechanisms for lexical tone and intonation, but did not provide any evidence to support the claim. There are some evidence of differential laryngeal muscular activity on electromyography for F0 changes at the beginning of a sentence and rapid F0
changes within an utterance in non-tonal language (Hong, Kim & Kim, 2001), which may suggest separate physiological control for long and short range F0 changes (i.e., intonation and tone) in tone languages. Although there might have been separate physiological mechanism for tone and intonation, the sharing of one dominant characteristic (i.e., F0) between them, might lead to compromises in the actual realization in speech production due to acoustic and articulatory constraints (Liu & Xu, 2005). Ma, Ciocca and Whitehill (2006; 2011) found that productions of tone and intonation in Cantonese are affected bi-directionally due to the simultaneous transmission of tone and intonation in the same speech signal. For example, while the general downstep of a statement has no impact on the F0 contour of the tone, the final-rise in the F0 contour of a question was shown to change the canonical form of all six tones. Additionally, a high-rising tone at the final position of a statement might alter the listeners' perception and misled them to perceive a statement as a question. This suggested that tone production in speakers with PD is likely to be affected, given monotone is one of the main deviant characteristics of PD speech.

In Cantonese speakers with PD, monotone speech is one of the key perceptual characteristics (Whitehill, Ma & Lee, 2003). In studying the question-statement contrast in Cantonese, studies showed that speakers with PD marked the intonation of questions in a similar manner to that of control speakers, but they were less efficient in exploiting some of the acoustic cues for marking questions (e.g. final F0 rise and increase in F0 level) (Ma & Whitehill, 2008; Ma, Whitehill & So, 2010). That is, the speakers with PD retained the phonological distinction in marking questions and statements, but the physiological constraints of PD led to a reduction in the F0 excursions at a phonetic level for some of the speakers. The reduction in intonation marking caused perceptual confusion among native listeners. Cantonese PD speakers have also been reported to have relatively intact lexical tone production (Whitehill, Ma & Lee, 2003; Ma, 2009). Although the reduction in F0 range of speakers with PD results in a smaller tonal space than control speakers, speakers with PD preserved similar F0 contours and tonal contrasts as control speakers (Ma, 2009). The largest difference between speakers with PD and control speakers in lexical tone production
occurred at the final position of questions (Ma, 2009). In non-dysarthric speakers, the canonical form of all lexical tones at the final position of question is modified to a rising contour as a result of F0 final-rise in questions. However, the magnitude of this modification, i.e., the magnitude of the interaction between intonation and tone, was reduced in speakers with PD (Ma, 2009).

Although both intonation and tone have F0 as their primary acoustic cue, findings from Cantonese speakers with PD show that it is possible to have different degrees of impairment in tone and intonation production. This supports the hypothesis of differential control between intonation and tone and serves as an example to caution researchers and clinicians against generalizing the findings from one language to another without consideration of the specific features of each language.

Conclusion

In this chapter, we reviewed the current findings for the assessment, treatment and characterization of motor speech disorders in Cantonese and Mandarin. We focused primarily on publications in English and a few easily accessible published studies in Chinese. We noted that there have been few published studies of apraxia in Chinese speakers. This review highlighted that the many of the speech characteristics of motor speech disorders in Chinese are similar to those reported for English-speaking individuals. This reflects the fact that motor speech disorders such as dysarthria are a direct consequence of neurophysiological impairment, which impact individuals similarly regardless of language. However, there are also some interactions between the neurophysiological impairment and the linguistic features of Chinese, such as tones. Studies have shown that is tone severely impaired in some speakers with dysarthric (e.g., with cerebral palsy), while relatively intact in some others (e.g., speakers with hypokinetic dysarthria associated with PD). Research in motor speech disorders in Chinese is still limited in many areas, especially in the lack of standardized assessment tools and the limited number of treatment studies. Finally, the diversity of languages in Chinese-speaking regions provides additional challenges to the study of motor speech disorders in Chinese.
References


dysarthria. *Journal of Speech and Hearing Research* 12, 246-269.


