Spectral moment analysis of affricates produced by Mandarin-speaking pre-adolescents with repaired cleft palate

Chenghui Jiang [a,b,*], Tara L. Whitehill [b,1], Bradley McPherson [b], Manwa L. Ng [b]

[a] Stomatological College, Nanjing Medical University, China
[b] Division of Speech and Hearing Sciences, Faculty of Education, University of Hong Kong, China

ABSTRACT

Objective: To explore the spectral differences in frication noise between aspirated and unaspirated affricates in typical Putonghua (standard Mandarin Chinese) pre-adolescent speakers, and to compare the spectral characteristics of affricate production between speakers with repaired cleft palate and their non-cleft peers.
Participants and intervention: Spectral moment analysis, a quantitative approach to capture the contour of speech spectra, was carried out on speech samples produced by two groups of speakers: (a) speakers with repaired cleft palate ($n = 14$, mean age = 11.7 years) and (b) typical speakers ($n = 10$, mean age = 11.0 years).
Results: Data from typical speakers showed that the unaspirated affricates had significantly higher first spectral moment (M1) than their aspirated counterparts. Compared with typical speakers, individuals with repaired cleft palate exhibited a lower first moment for the four affricates /ts, tʂ, tɕʰ, tɕ/.
Conclusion: The results revealed important acoustical differences between aspirated and unaspirated affricates for typical speakers. The trend of spectral deviation may have contributed to the difficulty in producing unaspirated affricates found in Putonghua-speaking individuals with speech disorders related to cleft palate.

© 2016 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Cleft palate is a congenital deficiency which may cause multiple problems including difficulties in speech, hearing, feeding and maxillofacial development. The speech problems associated with cleft palate can be characterized by phonological disorder and/or misarticulation, abnormal resonance (hypernasality and/or hyponasality) and nasal emission [1]. Studies of English, Cantonese and Putonghua speakers with repaired cleft palate demonstrated that affricates were more vulnerable to misarticulation than other manners of articulation [1–5]. The low accuracy and high variability associated with affricate production in speakers with repaired cleft palate prompted further exploration of the articulation of Putonghua affricates in the present research project, particularly as more affricates exist in Putonghua phonology than English [6].

An affricate is a combination of stop and fricative articulatory components [7] and it involves a period of complete occlusion of the vocal tract followed by a period of frication. Compared with English, more affricates exist in Putonghua phonology. In Putonghua, there are three pairs of affricates distributed at different articulatory places – the alveolar, post-alveolar (retroflex) and alveolar-palatal (palatal) regions. For each pair, the two affricates are distinguished from each other by aspiration; i.e., aspirated versus unaspirated. Similar to the voicing contrast in English and other languages, the aspiration contrast is manifested as a difference in voice onset time (VOT) [8]. The aspirated affricate has a longer VOT than its unaspirated counterpart [9]. Perceptual experiments on consonant accuracy associated with Mandarin speakers with repaired cleft palate indicated that for the group of speakers with perceived hypernasal resonance, the unaspirated affricates were more difficult to produce than their aspirated counterparts [2]. When producing alveolar affricates (/tsʰ/ and /ts/), a closure is firstly formed by elevating the tongue tip or blade to the alveolar ridge. Then a slight separation is made between the articulators to form the constriction for the fricative portion, resulting in a fricative noise with an outward airflow. For the retroflex pair (/tʂ/ and /tʂʰ/), they are typically described as apical-alveolar sounds made

by curling the tongue tip upward, and approaching the post-alveolar region using the underside of the tongue tip [10]. The major difference between the alveolar and the retroflex affricates is that the tip of the tongue is behind the front teeth for the former and behind the alveolar ridge for the latter [10]. Another unique group of affricates in Putonghua are the palatal pair, /tɕ/ and /tɕʰ/, which are similar to the palato-alveolar pair (/tʃ/ and /dʒ/) in terms of placement of articulation. However, the blade and the front of the tongue are raised to a much higher position [11]. Unlike their articulatory gestures, the acoustical features of Putonghua affricates have not been closely investigated.

Spectra of stop bursts and fricative noises provide valuable information regarding the articulatory configuration of consonants [10,12,13]. However, quantitative description of spectral shape can be laborious due to the complexity associated with spectra [14]. Therefore, Forrest et al. [12] developed a simple quantitative approach to measure and describe the spectral shape, namely spectral moment analysis (SMA). With SMA, the aperiodic energy is treated as a random probability distribution from which the first four moments can be derived [12]. The first (M1) and the second moment (M2) reflect the average energy and range of an energy spectrum, respectively. The third moment (L3, the normalized third moment, skewness) indicates the symmetry of a distribution. A distribution which is skewed to the right will have a positive skewness, and vice versa. The fourth spectral moment (L4, the normalized fourth moment) indicates the peakedness of the distribution.

The spectral features of normally articulated consonants have been explored in English speakers [12,13]. Forrest et al. [12] calculated the spectral moment features of three voiceless stops (/p, t, k/) produced by adult speakers. Their data (through linear transformed spectra) suggested that labial and alveolar stops were distinctive in terms of mean (M2) and skewness (L3). The velar stops are distinguished from the other two places by kurtosis (L4). Later, Jongman et al. extended Forrest et al.'s [13] study to investigate fricatives of four different places (/f, v; θ, ð; s, z; ʃ, ʒ/) and found that /s, z/ had the greatest spectral mean (M1) and kurtosis (L4), implying a greater average energy and high peakedness. /ʃ, ʒ/ showed the lowest spectral mean (M1) but consistently positive skewness (L3). This indicated a low average energy of the spectra, with the majority of energy distributed in the lower frequency region. Variance (M2) was low for the sibilants (/s, z; ʃ, ʒ/) and high for the non-sibilants (/f, v; θ, ð/). Several pilot studies have quantified the spectral properties of consonants produced by Putonghua speakers. Mays and Beckman [15] explored this issue using three unaspirated affricates (/ts, /tʂ, tɕ) produced by Putonghua-speaking men. Among the three places, / ts/ showed the highest spectral mean value (M1) and lowest skewness (L3). The kurtosis was the greatest for the retroflex /tʂ/. Lai [16] compared the spectral mean (M1) between retroflex (/tʂʰ, tʂ, ʂ/) and non-retroflex (/tsʰ, ts, s/) productions from a group of Taiwan Mandarin speakers. She reported that for the female speakers, the non-retroflex consonants had a much higher M1 than the retroflex. A study was conducted by Jiang et al. [17] to explore the relationship between the spectral moments features and listener perception of accuracy of a number of typical alveolar and retroflex affricate articulation from Mandarin pre-adolescents speakers. It was reported that the third moment (L3) showed a significant relationship with listener perceptual judgment of both typical alveolar and retroflex consonants. Lee [18] calculated the spectral mean values (M1) for three different fricative places (/s, ʂ, ɕ/) produced by six speakers. A strict ordering was found for the three sibilants in the first spectral moment: /s/ highest, /ʂ/ lowest and /ɕ/ intermediate. Similar to findings from English speakers, alveolar fricatives/affricates had the highest spectral mean (M1) in both languages (English and Putonghua) but there were different

findings for skewness (L3) and kurtosis (L4) values. In sum, spectral moment analysis may provide useful information on the distinctive place of articulation for the phonemes in a typical speaker's phonological inventory. However, few investigations have been conducted about the distinctive feature of aspiration.

Spectral moment analysis also shows some promise in capturing the spectral characteristics of consonants produced by individuals with motor speech disorders. Tjaden and Turner [19] investigated spectral moment features of two fricatives (/s/ and /ʃ/) in a group of seven speakers with amyotrophic lateral sclerosis (ALS). The difference between the two fricatives in spectral mean (M1) and skewness (L3) was less for speakers with ALS than for typical speakers. Similar findings of decreased spectral distinction between /s/ and /ʃ/ were also identified among elderly speakers with Parkinson disease [20]. The spectral distinctions between the two fricatives were considered to be correlated with perceived precision in consonant production. Smaller spectral distinctions indicated worsened perceived articulatory precision. Recently, two studies have used spectral moment analysis to describe the first moment characteristics of stops produced by children with repaired cleft palate who had maxillary arch anomalies (e.g., maxillary collapse) in two languages (English and Persian) [21,22]. For both languages, a significantly smaller value was found in the first moment (M1) of the alveolar stop /t/ but not for the velar stop /k/. The difference in the first moment (M1) between the alveolar stop and velar stop (/t/-/k/) was also significantly decreased. It was concluded that a decrease in the first moment (M1) could effectively signify a more posterior articulation of alveolar stops when maxillary arch anomalies occurred. Karlsson et al. [22] also found that the first spectral moment was sensitive to change in place of articulation on children with speech delayed. They reported that there was a decrease in M1 for palatalized sibilants and an increased M1 for a dentalized sibilant. Jiang et al. [17] also found that, different from typical speech, for productions by Mandarin-speaking pre-adolescents with cleft palate, the M1 was significantly correlated with listener perceptual visual analog scaling of the accuracy of retroflex affricates, but not for the alveolar affricates. Based on the above findings from speakers with speech disorders, it was suggested that spectral moment analysis could also identify deviations in articulation. Such misarticulations may cause distortion to listener perception of speakers with cleft disorders.

Based on the above findings, SMA is a potential tool for examining affricate production by speakers with repaired cleft palate. Therefore, the purpose of the present study was to: (1) explore the difference in spectral properties between the aspirated and unaspirated affricates of typical Mandarin speakers; and (2) systematically describe the spectral properties of all six affricates in Mandarin phonological system produced by Mandarin speakers with repaired cleft palate, as compared to typical speakers.

2. Method

2.1. Participants

Speech samples were excerpted from the data pool of a larger scale study (unpublished data) in which speech produced by 14 Mandarin speakers with non-syndromic cleft palate was examined. The cleft participant group was comprised of 12 males and 2 females, with ages from 10 to 14 years (mean = 11.7 years, SD = 1.3 years) (Appendix I). Of the cleft participants, five had unilateral cleft lip and palate and nine received bilateral cleft lip and palate. The participants were recruited at the Cleft Lip and Palate Centers of Jiangsu Stomatological Hospital and Xuzhou First People's Hospital, in mainland China. All participants were native Mandarin speakers and received primary palatoplasty earlier than