Journal of Phonetics 52 (2015) 26–34

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/phonetics

Research Article

Relative contributions of vowels and consonants in recognizing isolated Mandarin words

Fei Chen [a,*], Michelle L.Y. Wong [b], Shufeng Zhu [a], Lena L.N. Wong [b]

[a] Department of Electrical and Electronic Engineering, South University of Science and Technology of China, Shenzhen, China
[b] Division of Speech and Hearing Sciences, The University of Hong Kong, Hong Kong

ARTICLE INFO

Article history:
Received 23 April 2014
Received in revised form
11 April 2015
Accepted 16 April 2015

Keywords:
Mandarin word recognition
Lexical tone identification
Vowel–consonant transition
Noise-replacement paradigm

ABSTRACT

This study investigated the relative contributions of vowels and consonants in recognizing isolated Mandarin words. Normal-hearing native-Mandarin listeners were instructed to recognize isolated Mandarin words and identify consonants, vowels and tones with stimuli synthesized to contain different proportions of consonant or vowel segments, including five conditions of consonant-only, vowel-only, consonant or vowel plus consonant–vowel (C–V) transition, and C–V transition. The recognition score of the vowel-only Mandarin words was significantly higher than that of the consonant-only words; and word recognition scores had a higher correlation with vowel identification scores than consonant identification scores. Moreover, adding a small portion of C–V transition significantly improved the recognition score of the consonant-only Mandarin words. In the conditions of C–V transition and consonant plus C–V transition, the duration of preserved portion predicted modestly well the scores of isolated word recognition, and vowel, consonant and lexical tone identification in Mandarin. These findings suggest that there is a greater contribution of vowels than consonants to isolated word recognition in Mandarin, which is different from previous outcomes in English.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Vowels (V) and consonants (C) are two traditional categories of speech sounds, which are used to produce words in languages throughout the world (Ladefoged & Disner, 2012). Vowels are produced with a more open vocal tract, and their duration is longer. In contrast, the production of consonants involves constriction in the vocal tract, and their duration is shorter (Abercrombie, 1967; Wright, 2004). Due to their difference in the manner of production, vowels and consonants carry different acoustic information (Kent & Read, 2002). Formant pattern is the acoustic cue that is most important for identifying vowels. For consonants, the most relevant acoustic cues differ in terms of place and manner of articulation. In general, burst and formant transition are the acoustic cues for the place of articulation, while spectral pattern is the acoustic cue for the manner of articulation (Pickett, 1999).

Many studies have suggested that vowels and consonants have different roles in speech perception. For example, Nespor, Pena, and Mehler (2003) and Toro, Nespor, Mehler, and Bonatti (2008) found that vowels contain information about prosody which is used to interpret the syntactic structure, whereas consonants' higher distinctive power within a word provides more cues for lexical identification than vowels. There are a number of investigations on the relative contributions of vowels and consonants to speech (word and sentence) intelligibility (e.g., Chen & Loizou, 2012; Chen, Wong, & Wong, 2013; Cole, Yan, Mak, Fanty, & Bailey, 1996; Fogerty & Humes, 2010; Fogerty & Humes, 2012; Kewley-Port, Burkle, & Lee, 2007; Owren & Cardillo, 2006). The study of Cole et al. (1996) showed a 2:1 advantage of vowels over consonants in word recognition at the sentence level using noise-replaced stimuli (i.e., vowel-only sentences with consonants replaced by noise, and vice versa). Similar results were also found in Kewley-Port et al. (2007) and Fogerty and Humes (2012). A recent study concerning the segmental contribution to sentence intelligibility in Mandarin revealed a 3:1 advantage for the perception of the V-only sentences over the C-only sentences (Chen et al., 2013). This ratio was higher than that in English, suggesting that vowels might play an even greater role in Mandarin sentence intelligibility than consonants.

Speech perception of sentences is different from that of isolated words. First, the linguistic context in sentences helps listeners to predict the words by top-down processing in addition to acoustic cues (e.g., Chen, Wong, & Hu, 2014), whereas listeners mainly

* Correspondence to: Department of Electrical and Electronic Engineering, South University of Science and Technology of China, Xueyuan Road 1088#, Xili, Nanshan District, Shenzhen, China. Tel.: +86 755 88015878.
E-mail address: fchen@sustc.edu.cn (F. Chen).

0095-4470/$ - see front matter © 2015 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.wocn.2015.04.003

recognize isolated words by their acoustic features using bottom-up processing (Denes & Pinson, 1993; Kewley-Port et al., 2007). Second, there are prosodic changes in sentences (Shattuck-Hufnagel, & Turk, 1996). Since vowels are the main carrier of prosodic information and thus convey information about syntax (Nespor et al., 2003), the vowel advantage in word recognition at the sentence level may be related to the syntactic content itself. The relative contributions of vowels and consonants to sentence intelligibility may not be the same as those to isolated word recognition.

There are a few studies that examined the relative contributions of vowels and consonants to the perception of isolated English words (Fogerty & Humes, 2010; Owren & Cardillo, 2006). Owren and Cardillo (2006) showed that listeners discriminated talker identity better with vowel-only isolated words, and discriminated word meaning better with consonant-only isolated words using a forced-choice judgement task. Fogerty and Humes (2010) used a word identification task to investigate the segmental contribution to recognize isolated English words. Although no significant difference was found between the consonant-only and vowel-only words in word recognition performance, the performance with the consonant-only words was more influenced by lexical difficulty than that for vowel-only words, suggesting that consonants had a more important role in lexical access. The possible reasons for the greater contribution of consonants to lexical access included the larger number of consonants and more distinctive nature of consonants compared to vowels in English (Fogerty & Humes, 2010).

Although there are studies concerning the segmental contribution to isolated word intelligibility in English, to date there is no study systematically investigating the relative contributions of vowels and consonants to recognize isolated Mandarin words. Mandarin phonology is different from that in English. Mandarin is a tonal language with four lexical tones (i.e., high level, rising, falling and rising, and falling), while English is a non-tonal language (Howie, 1976). Moreover, Mandarin has a simpler syllable structure than English. Consonant clusters can appear in onset and coda position for English syllables (McMahon, 2002) but not for Mandarin. The numbers of vowels and consonants in Mandarin are also different from those in English. There are a greater number of vowels than consonants in Mandarin (i.e., 35 vowels vs. 21 consonants; Yin & Felley, 1990). In contrast, there are fewer vowels than consonants in English (20 vowels vs. 24 consonants; Jones, 2006). Therefore, the segmental contribution to the isolated word recognition in English may be different from that in Mandarin due in part to the acoustic, phonetic, and statistical differences between the two different languages.

Apart from the segmental contribution to isolated word recognition, it is also demonstrated that adding a portion of the consonant–vowel (C–V) transition (across the C–V boundary) improved isolated word recognition in English (Fogerty & Humes, 2010). Similar findings were obtained in the study of Mandarin sentences (Chen et al., 2013). It is believed that the C–V transition contains co-articulatory information which helps the identification of adjacent phonemes (Recasens, 1999). However, no study has investigated the contribution of the C–V transitions to the recognition of isolated Mandarin words.

The purpose of the present study is (1) to determine the relative contributions of vowels and consonants in recognizing isolated Mandarin words; (2) to compare the relative contributions of vowels and consonants in isolated word recognition between Mandarin and English; and (3) to assess the contribution of C–V transition in recognizing isolated Mandarin words. In order to find out possible reasons for the relative contributions of vowels and consonants in recognizing isolated Mandarin words, the relative contributions of vowels and consonants in identifying consonants, vowels and lexical tones were also examined in this study.

2. Method

2.1. Subjects and materials

Nineteen (8 males and 11 females, age range 18–27) normal-hearing native-Mandarin participants were recruited from The University of Hong Kong. All of them were undergraduates or postgraduates studying at The University of Hong Kong and came from mainland China. They all passed a hearing screening, and their bilateral pure tone thresholds were all below 20 dB HL at octave intervals from 250 to 8000 Hz (ANSI, 1996).

The isolated Mandarin words used were taken from a database of 1128 isolated Mandarin (monosyllabic) words, covering almost all daily-used words in Mandarin Chinese. All the words were spoken in isolation by a female native-Mandarin talker at a normal speaking rate and with broadcaster's voice quality. The fundamental frequency of recorded words ranged from 130 to 330 Hz.

The C–V boundaries (defined based on traditional segmental boundaries) were labeled manually by an experienced phonetician, and later verified by another experienced phonetician. All final nasal consonants were counted as part of their preceding vowels. The average duration of the words was 468 ms (consonants: 118 ms, vowels: 350 ms). The words include 21 consonants ([p], [pʰ], [m], [f], [t], [tʰ], [n], [l], [k], [kʰ], [x], [tɕ], [tɕʰ], [ɕ], [tʂ], [tʂʰ], [ʂ], [ʐ], [ts], [tsʰ], [s]), 35 vowels ([a], [o], [ɤ], [i], [u], [y], [aɪ], [eɪ], [ɑʊ], [oʊ], [ia], [iɛ], [iaʊ], [iəʊ], [ua], [uo], [uaɪ], [ueɪ], [yɛ], [an], [ən], [ɑŋ], [əŋ], [oŋ], [iɛn], [in], [iɑŋ], [iŋ], [iʊŋ], [uan], [uən], [uɑŋ], [ʊəŋ], [yɛn], [yn]), and 4 tones (high level, rising, falling and rising, and falling) (Yin & Felley, 1990).[1] The word database and vowel/consonant boundary information will be provided upon request. Note that the recordings of isolated Mandarin CV syllables yielded much longer vowel durations than if they were recorded in a sentence context, i.e., 350 ms (in this study) vs. 190 ms (Chen et al., 2013).

[1] In terms of Mandarin vowels, the 35 vowels in this study consist of 6 simple vowels, 13 complex vowels, and 16 compound nasal vowels (Yin & Felley, 1990). The vowel nuclei with a final nasal were grouped into vowels as many earlier studies classified them as part of vowels in Mandarin (e.g., Yin & Felley, 1990; Fu, Zeng, Shannon, & Soli 1998; Zhao, Patrick Rau, & Yang, 2005; Fu, Zhu, & Wang,, 2011).

Download English Version:

https://daneshyari.com/en/article/7532905

Download Persian Version:

https://daneshyari.com/article/7532905

Daneshyari.com