2pSCb24. Influence of duration on the perception of consonants /x/ and /j/ in Chinese

Li Feng*, Yoshitaka Nakajima and Ueda Kazuo

*Corresponding author's address: Kyushu University, 4-9-1 Shiobaru Minami-ku, Fukuoka, Japan, Fukuoka, 815-0033, Fukuoka, Japan, rhou1986@yahoo.co.jp

It was revealed that Mandarin Chinese native speakers' perception of /x/ and /j/ was affected by the duration of consonant parts. Two perceptual experiments, in which the method of constant stimuli was employed, were conducted. All participants were native speakers of Mandarin Chinese. In the first experiment, in which six MC native speakers participated, Chinese syllables that begin with /x/ or /j/ were extracted from a speech database, and the consonant parts were manipulated in terms of duration. As the duration of /x/ was decreased or the duration of /j/ was increased, to a certain extent, the consonant which had been originally /x/ was perceived as /j/, and vice versa. Synthesized noises instead of recorded consonants were utilized in the second experiment, in which two native speakers participated, similar effects of the consonant duration appeared.
1. Introduction

The present paper focuses on how Chinese-native listeners perceive two palatals in Mandarin Chinese, /x/ and /j/. These sounds should be formed with the back and middle of the tongue close to or touching the roof of the mouth to pronounce.

In a preliminary study, we investigated the syllable /xiao/ of Mandarin Chinese by manipulating the consonant part in terms of duration, and found an interesting phenomenon. When the duration of the consonant part /x/ was decreased to a certain extent, the syllable which had been originally heard as /xiao/ was perceived as /jiao/. Then we took up another syllable /jiao/ of Mandarin Chinese, and increased the duration of the consonant part. We found that the syllable which had been originally heard as /jiao/ was perceived as /xiao/. We considered that our perception of the consonant /x/ and /j/ could be influenced by the duration of the consonant part, and conducted some preliminary experiments. The present study was planned to confirm the results.

2. Experiment 1

To investigate how perception of /x/ and /j/ was affected by duration of the consonant part, we conducted an experiment utilizing stimuli beginning with a consonant part that was originally /x/.

2.1 Method

2.1.1 Stimuli

We extracted syllables that begin with /x/ from sentences spoken by two males and two females’ speakers in a speech database (NTT-AT Multilingual Speech Database 2002). We manipulated the consonant parts in terms of duration. We decreased the duration of the consonant part in steps of 10 ms. We always took a rise time of 10 ms at the beginning. There was also a condition in which only the rise time was shaped. The original syllables were also used.

2.1.2 Procedure

The auditory stimuli were generated by a personal computer (Dell D531). They were presented diotically through headphones (Stax SR-303) at 70 dBA (Fast-peak), measured with an artificial ear (Brüel & Kjær Type 4153).

The participant was instructed to watch the computer screen. Once the start button was clicked by the participant, a stimulus was presented three times. The order of the stimuli was randomized. Each experimental block was preceded by a warm-up
period of four trials. The participants were instructed to write down the syllables they heard on a sheet of paper. When a syllable of Mandarin Chinese was clearly perceived, they wrote down the syllable; when several different syllables were heard, they wrote down all the syllables from the most dominant one; when only a part of a syllable was heard, they wrote down the vowel or the consonant they perceived; when no speech sound was heard, the participant wrote “Did not hear.”

2.2 Results

![Result of subject 1](image1)

**FIGURE 1.** Result of Experiment 1. The x-axis shows the duration of the consonant part and the y-axis the percentage of the perception of /x/ or /j/.

FIGURE 1 about the results. With the decrease of the consonant-part duration, the perception of /x/ decreased, and the perception of /j/ increased. With a consonant-part duration longer than 100 ms, the perception of /x/ took place 100%, but when the consonant duration was 60-100 ms, the percentage dropped suddenly, and the perception on /j/ suddenly increased. With consonant-part duration shorter than 50 ms, the perception of /j/ took place 100%.

2.3 Discussion

The results showed that with the decrease of the consonant-part duration, to a certain extent, the consonant which had been originally /x/ changed into /j/. However, with the change of the consonant-part duration, other physical factors also must have changed. At the beginning of the consonant part, there may be some auditory cue facilitating the perception of /x/. When we decrease the duration of the consonant part in steps of 10 ms, this cue may be discarded. Then we cannot say that the change of consonant perception was caused by the duration difference. It was thus necessary to conduct another perceptual experiment utilizing stimuli in which the consonant parts were made by synthesized noise.
3. Experiment 2

To investigate whether similar effects of the consonant duration would appear with stimuli of synthesized noises and vowels instead of recorded speech, an experiment was conducted.

3.1 Method

3.1.1 Stimuli

We observed various Chinese syllables that began with /x/ and /j/ in the database, and found that the average duration of /x/ was about 120 ms, and that the average duration of /j/ was about 70 ms. Synthesized noises with the duration of 120 ms and 70 ms were thus utilized for these consonants. We synthesized vowel parts as harmonic complex tones. We analyzed vowels’ formants in Mandarin Chinese, and synthesized the vowel /i/ with F1=300, F2=2500, and F3=3300, and while /ü/ with F1=250, F2=1750, and F3=2200, we then connected the formants to the synthesized noises with an overlapping duration of 10 ms. As Mandarin Chinese is a tone language, we used 4 types of fundamental frequency patterns (170-160 Hz for tone 1, 130-170 Hz for tone 2, 160-120-180 Hz for tone 3, and 200-130 Hz for tone 4). With two kinds of consonants, 2 kinds of vowels, 2 kinds of energy changes and 4 kinds of tones, there were 10 experimental conditions.

3.1.2 Procedure

The signals were generated with a computer (Frontier) and an audio processor (ONKYO, SE-U55GX), and delivered to an amplifier (Stax SRM-313). The auditory stimuli were presented diotically through headphones (Stax SR-303) with the sound pressure at 65 dBA (Fast-peak). The levels were measured with an artificial ear (Brüel & Kjær, Type 4153).

The participant was instructed to watch the computer screen. Once the start button was clicked by the participant, the stimulus was presented 3 times. The order of the stimuli was randomized. The participants were instructed to write down the syllable they perceived on a sheet of paper.

Two normal-hearing native speakers of Mandarin Chinese, who were students of Kyushu University, participated.
3.2 Results

FIGURE 2. Intelligibility scores of the synthesized syllables. The difference between /x/ and /j/ before the same vowel was only the difference in the consonant-part duration: 120 or 70 ms. Different vowels had different formant frequencies.

FIGURE 2 illustrates the results. The intelligibility scores of the consonants were quite high. This means that, just with the different duration of the synthesized noises to be perceived as consonants, different consonants /x/ and /j/ were perceived clearly. The results supported our prediction that Mandarin Chinese native speakers’ perception of /x/ and /j/ was strongly affected by the duration of the consonant parts.

As Japanese language have consonants similar to /x/ and /j/ in Chinese, the same phenomenon may appear in Japanese, and we are planning to conduct an experiment utilizing Japanese stimuli.

References
1. Aihara, S. (2011) From the foundation of pronunciation to learn Chinese, Asahi, Tokyo