SNAP TO READ

PRONUNCIATION CHALLENGES OF MANDARIN PHONEMES ZH, CH, SH, D, T AMONG SMPIT CHENG HOO MAKASSAR STUDENTS

Nerli utami, Jufri, Nirdayanti Mustarih
Universitas Negeri Makassar
nerlyutami@gmail.com, jufriandipalewai@gmail.com, nirdayanti@unm.ac.id

First received: September 24, 2025 Final proof received: November 20, 2025

Abstract:
The differences in the phonological systems between Indonesian and Mandarin often cause difficulties for learners, especially in the consonant sounds zh, ch, sh, d, and t which do not have direct equivalents in Indonesian. These pronunciation difficulties, especially in junior high school students, are not free from errors in pronouncing the consonants h, ch, sh, d, and t, so this study shows and describes the form of pronunciation difficulties of the consonants zh, ch, sh, d, and t directly in grade IX students of SMP IT Cheng Hoo Makassar, and identifies the factors causing these difficulties. The research method used in this study is descriptive qualitative with 15 students as subjects. Data were collected through pronunciation tests, syllable tests, observations, structured interviews, and audio recordings. The results showed that most students had difficulty pronouncing the retroflex consonants zh and ch, as well as the aspirated consonant t. Errors that emerged included difficulties in tongue articulation, failure to distinguish between aspirated and non-aspirated sounds, and interference from Indonesian, for example, zh is pronounced as j and t is pronounced as d. Physiological factors such as a stiff tongue position also affect pronunciation accuracy. In conclusion, ninth-grade students of SMP IT Cheng Hoo Makassar still face significant challenges in pronouncing certain consonants in Mandarin. These

findings emphasize the need for more targeted phonetic teaching strategies, such as explicit articulation exercises and the use of audio-visual media, so that students can improve their accuracy and confidence in speaking Mandarin.

Keywords: *Pronunciation difficulties, language interference, consonant phonemes, mandarin*

INTRODUCTION

Language plays a crucial role in human communication, enabling individuals to convey ideas, emotions, and information. Mandarin is gaining popularity in Indonesia as economic, cultural, and educational cooperation between Indonesia and China increases. As one of the most widely spoken international languages, Mandarin is now taught at various levels of education, including junior high school. Mandarin has even become a foreign language taught in Indonesian schools, particularly in Makassar. Teaching Mandarin is not only seen as an additional skill but also as an effort to equip students with global competencies from an early age. Schools that include Mandarin as a subject strive to integrate religious values with foreign language learning relevant to 21st-century needs. This is interesting to study, as it demonstrates the integration of character education and language skills.

However, pronunciation difficulties experienced by students can actually become obstacles to achieving these goals if not properly identified and addressed. In addition to theoretical and practical aspects, research on pronunciation difficulties in Mandarin phonemes also holds strategic value in the context of international relations. Based on this, the importance of good Mandarin language skills will be a crucial asset for the younger generation. Therefore, effective phonetics learning in schools is more needed to help students in academic contexts, and also prepare them to face the challenges of cross-cultural communication in the future.

Many Indonesian learners face difficulties in mastering the phonological aspects of Mandarin, especially in pronouncing consonants that do not exist in Indonesian. Some consonant sounds such as zh [ʈ͡ʂ],

ch [tʂʰ], sh [ʂ], d [t], and t [tʰ] are often pronounced incorrectly, causing miscommunication and lowering students' self-confidence. For example, the word 照相 (Zhàoxiàng) is often pronounced by students as Chaoxiang or Shaoxiang, and the word 明天 (Míngtiān) is often pronounced as Míngdiān. Mistakes like these occur because Indonesian does not recognize retroflex or aspirative sounds, so students tend to replace foreign sounds with more familiar sounds. Language interference, or negative transfer, occurs when the phonetic habits from a learner's first language (L1) influence the production of sounds in the second language (L2), resulting in difficulties in achieving native-like pronunciation. In the case of Indonesian learners of Mandarin, the absence of retroflex sounds (zh, ch, sh) and differences in aspiration (d, t) in Indonesian phonology often leads to substitution or mispronunciation of these sounds. This theoretical perspective helps to understand why certain pronunciation errors commonly occur among students. According to the theory of articulatory phonetics (Ladefoged, 2006), sounds that do not exist in the learner's first language will be more difficult to pronounce correctly. Contrastive analysis (Lado, 2007) also emphasizes that differences between the first language and the target language can lead to negative transfer or interference. In this context, Indonesian language learners tend to replace zh, ch, and sh with j, c, and s, and experience confusion in distinguishing aspirated and non-aspirated alveolar pops such as t and d.

Several previous studies have shown that retroflex and aspirative sounds are the biggest challenges for non-native Mandarin learners (Anggreani, Marsuki, Amanda, & Irawati, 2021), (Sugiarti, Kusumaningtyas, Phanata, Rudiansyah, & Lianisyah, 2022). Research by (Nurhusnah, Rahmanita, & Sakinah, 2024) also revealed that mother tongue significantly influences students' pronunciation, often leading to substitution errors. However, studies on the pronunciation difficulties of junior high school students in Indonesia, particularly related to these consonants, are still limited. Most previous studies have focused on analyzing consonants of the same type or within the same sound family, and were conducted with university students and adult

DOI: dx.doi.org/10.21274/ls.2025.17.2.349-378

learners. This focus differs from this study, which seeks to examine the level of pronunciation difficulties in junior high school students in more depth.

According to the theory of Articulation Phonetics, as explained by (Ladefoged, 2006), each language sound can be classified based on the place of articulation, manner of articulation, and voicing. If a sound does not exist in the learner's mother tongue, it will be more difficult for students to articulate it accurately. This difficulty is clarified through the Contrastive Analysis approach proposed by (Lado, 2007). In this approach, it is explained that the difference in sound structure between the first language (L1) and the second language (L2) will cause negative interference, where students tend to use the sound system of L1 when pronouncing sounds in L2. In this context, students of SMP IT Chengho Makassar who are native Indonesian speakers tend to replace the sounds zh, ch, sh with j, c, s because they are more familiar to the ear and easier to pronounce. Similarly, the sounds d and t in Mandarin are often pronounced without paying attention to the difference in aspiration, because in Indonesian, this difference is not a factor in differentiating meaning. Furthermore, according to the language interference theory explained by (Weirench, 1953), errors in foreign language pronunciation often occur due to the influence of the mother tongue. This interference can occur in the form of substitution, distortion, or omission of sounds.

In addition to linguistic factors, physiological aspects also influence pronunciation accuracy. For example, the shape of the tongue, teeth, or jaw can affect a student's ability to produce retroflex and aspirated sounds (Awaru, 2024). These factors are often overlooked, even though they can both hinder and support the process of acquiring foreign language phonology. In fact, the junior high school stage is a crucial phase in foreign language acquisition because during this period, students have high cognitive flexibility to absorb new sounds. By directing the study at this level, research can provide a more comprehensive picture of phonological barriers from the early stages of learning. The Mandarin phonological system has unique characteristics that

significantly differentiate it from Indonesian. One of the main differences lies in the presence of retroflex consonants (zh, ch, sh) and the aspirated-non-aspirated contrast (t and d). In Indonesian, differences in aspiration do not function as a difference in meaning, while in Mandarin, small errors in aspiration can result in different meanings. For example, the aspirated word 他 tā (he) has a different meaning than the unaspirated 大 dà (big). This phonological difference is one of the main reasons why Indonesian students often have difficulty pronouncing Mandarin sounds correctly.

Referring to the explanation above and previous research outlined, research related to pronunciation difficulties at the junior high school level is still limited, particularly schools in Makassar. Therefore, this study was conducted to analyze the pronunciation difficulties of the consonants zh, ch, sh, d, and t among ninth-grade students at SMP IT Cheng Hoo Makassar and identify the factors causing these difficulties. By addressing these issues, the research is expected to contribute to the design of more effective Mandarin phonetics learning strategies that can improve pronunciation accuracy and student confidence in Mandarin.

METHOD

This study uses a descriptive qualitative approach as explained by (Creswell, 2012) which states that "descriptive qualitative research is a research approach that aims to understand and describe in depth the phenomena experienced by individuals or groups in a natural context, without any manipulation of variables or intervention from the researcher", with the aim of exploring in depth the difficulties in pronouncing Mandarin consonants in ninth-grade students of SMP IT Cheng Hoo Makassar. All 15 ninth-grade students were selected as respondents for the study through a total sampling technique, because the population size was relatively small, allowing researchers to examine the entire population. Respondents consisted of male and female students aged 14–15 years, who had studied Mandarin for at least two years.

DOI: dx.doi.org/10.21274/ls.2025.17.2.349-378

Data collection was conducted in three stages: a pronunciation test, a syllable test, a structured interview, reading recording, and direct observations during testing. The pronunciation test was used to identify the extent to which students were able to pronounce the consonant sounds zh, ch, sh, d, and t correctly in sentence contexts. This test provided data on pronunciation performance in situations that resemble natural language use, thus addressing research questions related to students' difficulties at the word and sentence levels. Each of the five target consonants was embedded in 10 sentences, and the test was conducted twice a week for one month.

In addition, the syllable test was designed to assess the students' ability to distinguish phonetically similar sounds, such as zhā–chā or ti–di. The results helped answer research questions related to the perception and production of specific phonemes, as well as errors influenced by differences between the Indonesian and Mandarin sound systems. This test consisted of at least five minimal pairs: zhā – chā, shū – zhu, ti – di, tong – dong, che – zhe, which were also conducted twice a week for students to pronounce the syllables to find out the differences in their sounds.

Furthermore, structured interviews were conducted to obtain supporting information about students' perceptions of pronunciation challenges and their strategies in learning Mandarin sounds. Reading recordings served to analyze pronunciation errors in longer and more structured discourse. Finally, observations were used to document students' behaviors, difficulties, and strategies during each testing session, providing contextual insights that enriched the data analysis. Together, the qualitative data enriched the test findings by providing a deeper understanding of the contributing factors behind pronunciation errors and the students' level of phonological awareness.

In the structured interview phase, ten questions were asked, covering factors that contribute to students' difficulties in pronouncing Mandarin consonants. Interviews were conducted once in the third week to gain a deeper understanding of students' pronunciation difficulties. The recording stage

involves video recording the pronunciation of the syllables and consonants zh, ch, sh, d, and t by students as follows: zhā – chā, shū – zhu, ti – di, tong – dong, che – zhe. This phase provided additional data for researchers to more accurately observe the students' pronunciation patterns and pronunciations.

Each test was conducted individually, with important observation systematically recorded by the researcher throughout the process. Interviews were conducted after the tests to obtain more in-depth information regarding linguistic and non-linguistic factors influencing pronunciation difficulties. All collected data was systematically transcribed for subsequent analysis. Data analysis was conducted qualitatively by comparing students' pronunciation results with standard Mandarin pronunciation and categorizing errors into substitutions, omissions, and articulation errors. Furthermore, test and observation results were triangulated with interview data to obtain a more comprehensive picture of the factors causing pronunciation difficulties experienced by students.

To ensure data credibility in this study, three procedures were applied. First, method triangulation was conducted through the use of pronunciation tests, syllable tests, reading recordings, interviews, and observations to validate findings from multiple perspectives. Second, prolonged engagement and persistent observation were implemented by conducting repeated testing over one month to obtain stable and authentic pronunciation performance data. Third, the researchers consulted with a research lecturer with expertise in Mandarin phonology to review the analysis results to ensure accuracy of interpretation and reduce the potential for subjective bias. These procedures collectively ensure that the findings of this study are credible and reliable.

FINDINGS AND DISCUSSION

This section presents the results of a study on the pronunciation difficulties of the consonants zh, ch, sh, d, and t among ninth-grade students at SMP IT Cheng Hoo Makassar. The results are organized based on the research objectives and then discussed by linking them to theory and previous

DOI: dx.doi.org/10.21274/ls.2025.17.2.349-378

research. Therefore, the following description not only presents empirical findings but also provides an in-depth explanation of the causes and forms of pronunciation difficulties experienced by students.

Difficulties in the Pronunciation of Mandarin Consonants

Based on the results of observations and data analysis, it was found that most students had difficulty pronouncing the consonants zh, ch, and sh, as well as d and t in Mandarin.

Tabel 1. Difficulties in the Pronunciation of Mandarin Consonants

Category	SM	M	C	S	SS	Conclusion
Consonants zh	1	0	1	12	1	93,33%
Consonants ch	0	2	3	10	0	86,67%
Consonants sh	6	9	0	0	0	0,00%
Consonants d	9	5	0	1	0	6,67%
Consonants t	0	3	3	9	0	80,00%

a. Consonant zh

Out of the 15 samples, 12 students were found to have difficulty pronouncing the consonant zh. The zh consonant is a distinctive retroflex consonant. It is pronounced similarly to the English letter “j,” but with the tongue curled further back, and without a strong burst of air. Students tend to struggle with both articulating and distinguishing the zh sound due to insufficient tongue retroflexion. Additionally, some students produce an audible aspiration, resulting in a pronunciation that resembles the ch or sh sounds.

b. Consonant ch

Out of the 15 samples, 10 students were found to have difficulty, and 3 students reported moderate difficulty in pronouncing the consonant ch. The ch consonant is the aspirated counterpart of zh, and it is produced with a strong burst of air. Many students struggle with this sound due to the close similarity between zh and ch in both articulation and tongue placement. As a result, students often confuse the two sounds when pronouncing them.

c. Consonant sh

Among the consonants zh and ch, students found it easier to pronounce the consonant sh. This is because the tongue position for sh is placed further forward in the mouth, making it easier to articulate. Additionally, the airflow produced during the pronunciation of sh resembles a hissing sound, which helps students pronounce it more accurately.

d. Consonant d

Out of the 15 samples, students found it easier to pronounce the consonant d. This is because the sound of d closely resembles the Indonesian t sound as in the word "tidak," and it is a non-aspirated consonant, meaning it is produced without a strong burst of air.

e. Consonant t

Out of the 15 samples, 9 students were found to have difficulty pronouncing the consonant t. The t consonant in Mandarin is an aspirated sound, produced with a noticeable burst of air. However, students are accustomed to pronouncing t as in Indonesian, which is typically unaspirated (produced without a burst of air). This difference in aspiration often leads to mispronunciation.

Several tests were conducted to identify the various difficulties experienced by the students, including the following:

1. Pronunciation Errors in Syllables

To identify the difficulties faced by students in pronouncing the consonant sounds zh, ch, sh, d, and t, a syllable test was conducted. Each pair was designed as a minimal pair, differing by only one phonemic segment, with the aim of assessing students' ability to distinguish and produce the sounds.

Tabel 0.2. Pronunciation Errors in Syllables

Pair	Number of Students Experiencing Difficulty	Persentase
zhā – chā	6	40%
shū – zhū	0	-
tī – dī	0	-

DOI: dx.doi.org/10.21274/ls.2025.17.2.349-378

dōng- tōng	9	60%
chē- zhē	9	60%

Table 2. shows the test results showed that 60% of the students experienced major difficulties with pair 5 (che-zhe) and pair 4 (dong-tong). Meanwhile, 40% of the students had moderate difficulty distinguishing pair 1 (zha-cha) due to the high similarity in sound and nearly identical tongue placement in the deeper part of the mouth. This was also attributed to students still struggling with articulation—the clear pronunciation process that involves the speech organs (tongue, palate, lips). Students sometimes pronounced the consonant zh with unclear emphasis, causing it to sound like ch or sh. Most students performed quite well in distinguishing pair 2 (shu-zhu).

2. Pronunciation Errors in Word Forms

Tabel 0.3. Pronunciation Errors in Word Forms

Original Word	Transformed Word
Chángcháng	Cángcáng
Dà	Tà
Pāizhào	Pāichào

Table 3. presents the results pronunciation test revealed that many students made errors both in consonant sounds and in the articulation of words that did not conform to the phonetic rules of Mandarin. Many students experienced difficulties due to the similarity of consonant sounds and challenges in articulation (the use of speech organs). For example, some students pronounced consonants that should be aspirated (with a burst of air) as non-aspirated (without a burst of air), and some who were supposed to pronounce the consonant zh instead produced a sound resembling sh. However, aside from the consonants zh, ch, and t, students performed better in pronouncing sh and d. Many students reported that the sh consonant was easier to pronounce compared to zh, ch, and t, largely because the consonant d does not require aspiration, which makes it easier for students to articulate. Additionally, the pronunciation of d feels more natural since students do not

need to adjust their tongue or mouth placement extremely. Most pronunciation errors were caused by insufficient tongue curling. Nevertheless, not all consonants contributed equally to the errors.

3. Pronunciation Errors in Distinguishing the Consonants zh, ch, sh, d, and t

As many as 93.33% of students tended to have difficulty distinguishing between the sounds zh and ch, as they perceived these two consonants to be quite similar. Meanwhile, 6.67% of students had difficulty pronouncing the aspirated consonant t, which requires a burst of air. However, students found it easier to pronounce the consonant d because it does not require aspiration, thus causing less difficulty. Similarly, the consonant sh was easier to pronounce since it involves a continuous flow of air, making it less challenging for the students.

Factors Causing Pronunciation Difficulties

The findings from the pronunciation tests and interviews indicate that the difficulties are shaped by four main contributing factors: linguistic, physiological, psychological, and pedagogical.

a. Linguistic Factors

Students' pronunciation difficulties are mainly influenced by linguistic interference from their first language, Bahasa Indonesia. Bahasa Indonesia does not distinguish between retroflex and alveolar sounds, nor between aspirated and unaspirated consonants. "I don't really know the difference between zh and ch because they sound the same to me" (S3). This finding aligns with (Anggreani, Marsuki, Amanda, & Irawati, 2021)who assert that first-language interference significantly contributes to pronunciation errors among Indonesian learners of Mandarin. The limited phonological system of Bahasa Indonesia thus becomes a major obstacle in mastering unique Mandarin phonemes.

b. Physiological Factors

Some students experience difficulty positioning their tongue correctly when producing retroflex sounds. "My tongue feels stiff when I try to say

DOI: dx.doi.org/10.21274/ls.2025.17.2.349-378

zh," (S5). This statement indicates a lack of flexibility in the speech organs. As noted by Nurhusnah et al. (2024), physiological factors, especially control over the articulatory organs, play a vital role in successful pronunciation. Without sufficient mastery of articulation techniques, the sounds produced tend to resemble other, easier phonemes.

c. Psychological Factors

Emotional states such as nervousness and low self-confidence also affect pronunciation accuracy. "When I get nervous, I speak too fast and don't realize I'm saying it wrong," (S2). This shows that affective factors directly influence speech production. Sugiarti et al. (2022) emphasize that explicit articulation training can help students overcome such psychological barriers and improve control over their speech.

d. Pedagogical and Environmental Factors

Phonetic guidance from instructors is essential in helping students overcome pronunciation challenges, particularly when teachers provide consistent and accurate pronunciation models. "I can pronounce it better when the teacher shows me how to say it," (S7). Exposure to native speaker audio also enhances students' auditory discrimination of similar sounds. Furthermore, a supportive learning environment—both in and outside the classroom—positively contributes to improvement. Interview results reveal that most students benefited from repeated pronunciation drills and articulation practice, although some reported receiving little explicit instruction.

Indonesian Language Interference on the Consonant Sounds zh, ch, sh, d, and t

In examining the factors that contribute to students' pronunciation difficulties, one of the key challenges experienced by Indonesian learners of Mandarin is the interference of their native language, particularly in the articulation of specific consonant sounds. The sounds zh, ch, and sh do not exist in the Indonesian phonological system, while d and t are pronounced differently in terms of aspiration. This phonetic gap often leads

to mispronunciations, which can affect students' overall speaking accuracy and comprehension in Mandarin. Language interference, or negative transfer, occurs when the phonetic habits from the first language (L1) influence the production of sounds in the second language (L2), resulting in difficulties in achieving native-like pronunciation.

a. Different Articulation

Students who are native Indonesian speakers tend to simplify or substitute sounds with more familiar articulations, for example, replacing the consonant zh with c (e.g., Zhàoxiàng becoming Chaoxiang). The consonants zh and sh are pronounced with the tongue curled backward, a position that is uncommon in Indonesian.

b. Aspirated and Non-Aspirated

Indonesian does not distinguish aspiration, so all consonants such as t are pronounced without a strong burst of air. Indonesian native speakers often pronounce t and ch without aspiration, causing the sounds to appear less distinct or weak to target language speakers. For example, the consonant t may be pronounced as d (e.g., dong instead of tong).

c. Different Types of Consonants

In Mandarin, the consonants zh and sh produce friction sounds (fricatives), while d and t are stop consonants (plosives). However, Indonesian does not have the zh consonant and rarely uses sh naturally. As a result, students tend to confuse these sounds or pronounce them inconsistently due to unfamiliarity. Additionally, Indonesian does not have retroflex consonants; it only has alveolar consonants.

Difficulty Pronouncing the Consonants zh, ch, sh, d, and t

Differences in Phonological Systems

Indonesian does not have retroflex sounds such as zh, ch, and sh, so students tend to replace them with the closest similar sounds like j, c, or s. Mandarin has both aspirated and non-aspirated consonants, whereas Indonesian does not. According to (Sugiarti & phanata, 2022), the consonant sounds zh and ch do not exist in Indonesian and are quite difficult to

DOI: dx.doi.org/10.21274/ls.2025.17.2.349-378

pronounce, requiring precise technique by curling the tongue 90 degrees toward the palate. The zh consonant is non-aspirated, while ch is aspirated. These errors are consistent with the theory of phonological interference, where the mother tongue influences the pronunciation of a foreign language. This process is not easy and requires special practice and high phonetic awareness.

Without proper understanding and training, students will struggle to adapt tongue placement and articulation techniques needed to produce accurate retroflex sounds. Therefore, understanding these phonological differences is crucial in designing effective Mandarin teaching strategies. Teachers need to provide explanations and special exercises to help students recognize and distinguish these sounds, as well as guide them in mastering appropriate articulation techniques. The use of audio-visual media, direct articulation practice, and consistent correction can help minimize phonological interference from the mother tongue and accelerate the acquisition of new sounds.

Many students pronounce the word （照相, Zhàoxiàng）incorrectly according to Mandarin phonetic standards. Most students say Chaoxiang or Shaoxiang instead of Zhàoxiàng itself. Errors in consonant sounds and incorrect tongue placement often lead to mispronunciations, which can interfere with the clarity and meaning of the intended word. Similarly, with the word （明天, Míngtiān）, many students confuse the consonants t and d. The consonant t is an aspirated consonant that involves a burst of air, whereas d is a non-aspirated consonant (without a burst of air). Pronunciation errors occur because students are unable to properly identify and control aspiration, especially in the middle or final positions of syllables. When students pronounce a word, they often unconsciously neglect to control their articulation process—tongue position, mouth shape, and the adjustment between aspirated and non-aspirated sounds are overlooked. For example, in the word Míngtiān, the t consonant is often pronounced like d, which causes the word to sound phonetically incorrect. This is largely due to the influence of

Indonesian, which does not use aspiration as a meaning-distinguishing feature between /t/ and /d/. In Indonesian, the difference between t and d focuses more on voicing (voiced vs. voiceless) rather than the strength of air burst.

Besides errors in sound pairs such as zh–ch and d–t, pronunciation errors were also found in the word 常常 (chángcháng), which some students pronounce as cángcáng. In the word 常常 (chángcháng), the consonant ch is a voiceless aspirated retroflex consonant in Mandarin phonology. Producing this sound requires the speaker to curl the tongue towards the back of the palate while producing a strong burst of air. If the aspiration is absent or not clearly heard, the sound will resemble c, which is a non-aspirated consonant articulated more forward (dental-alveolar). Therefore, students tend to pronounce 常常 (chángcháng) as cángcáng, which phonologically indicates that the initial ch consonant is not pronounced with sufficient aspiration, making it sound like the unaspirated c consonant.

Articulation Difficulties

The sounds zh and ch require the tongue to be curled inward, a position that is not commonly used by native Indonesian speakers. Six students reported that their tongue felt "too curled," making it difficult to control their articulation. Articulation itself is a physical process of producing speech sounds that involves various speech organs, including the tongue, lips, teeth, palate, and vocal cords. In learning Mandarin, pronunciation errors frequently occur due to a number of factors, including interlinguistic and extralinguistic factors. Interlinguistic factors refer to the influence of the speaker's native language system, which differs from the language being learned or spoken. Extralinguistic factors refer to elements outside the language system that affect language use. This aligns with (Afrina & Celeveresty, 2020), who stated: "The analysis results show that the errors are classified as interlingual errors.

From the collected data, most students still apply their native language when learning Mandarin." Several phonetic processes contribute to these errors. One is labialization, the rounding of the lips during primary

DOI: dx.doi.org/10.21274/ls.2025.17.2.349-378

articulation, and another is retroflexion, which involves curling the tip of the tongue backward toward the palate during articulation. Because Indonesian speakers are not accustomed to positioning the tongue in a post-alveolar (retroflex) position, many students struggle to control tongue movement and muscular tension. Some students reported that their tongue felt "too curled" or placed unnaturally, leading to a loss of articulation accuracy and clarity in pronunciation.

This difficulty is further exacerbated by a lack of articulatory awareness—the ability to consciously control and monitor the position and movement of speech organs. Without sufficient phonetic training, students tend to substitute zh and ch with more familiar Indonesian sounds such as s, j, or c, which have different places of articulation and do not match the acoustic quality of the target sounds. This is supported by (liyanisah, sugiarti , & rudiansyah, 2022), who noted: "The analysis results show that the errors are classified as interlingual errors. These interlingual errors are caused by the interference of the mother tongue in the Mandarin language being learned.

Based on the data obtained, most students still apply their native language when learning Mandarin." To identify the difficulties faced by students in pronouncing the consonant sounds zh, ch, sh, d, and t, a syllable test was conducted. Each pair was designed as a minimal pair, differing only by one phonemic segment, to evaluate the students' ability to distinguish and produce each sound. Test results revealed that the greatest difficulties were found in Pair 5 (che–zhe) and Pair 4 (dong–tong), indicating key areas where students struggled to perceive and produce the contrasting features, particularly retroflexion and aspiration.

a. The pair zha-cha

This pair can be used to test whether students are able to distinguish between aspirated and non-aspirated retroflex sounds. These two sounds do not exist in the Indonesian language, causing students to tend to pronounce them with a similar tongue position and breath expulsion. As a result, the sound zha may be perceived as cha. This error indicates that students still

have difficulty differentiating between aspirated and non-aspirated sounds.

b. The pair shu-zhu

In this pair, all students were able to distinguish the sounds quite well, as both have closely related places of articulation, and errors were less frequent compared to the zha-cha pair. Students tended to pronounce the sound shu more accurately, not only because its tongue placement is easier, but also because the pronunciation of sh is relatively similar to sounds found in Indonesian.

c. The pair ti-dong

This pair is intended to assess sensitivity to vowel contrast, with 80% of students able to distinguish the two syllables accurately. Students tend to better understand ti with the aspirated feature because it consists of only two letters; however, when the syllable is longer, they tend to pronounce ti as non-aspirated. Nevertheless, 20% of the students still make errors, resulting in the consonant t being pronounced with aspiration (accompanied by a burst of air).

d. The pair dong-tong

This pair involves the contrast between d and t, and the results indicate that students often struggle to focus on distinguishing these two sounds, both in listening and pronunciation. Additionally, the aspect of articulation presents a significant challenge for students, as they must properly regulate the airflow pressure to produce the aspirated sound in t. Imperfections in distinguishing these pronunciations can potentially lead to communication inaccuracies and reduce speech clarity, which aligns with the findings of Supradi (2015). Mispronunciation of aspirated consonants as non-aspirated sounds is demonstrated by the difference in distinctive features related to the [+aspiration] characteristic.

e. The pair che-zhe

The error rate for this pair is relatively higher, similar to the zha-cha pair, as both involve aspirated and non-aspirated retroflex sounds. Based on the interview results, many students expressed that the consonant sounds

DOI: dx.doi.org/10.21274/ls.2025.17.2.349-378

che and zhe are very similar, yet they also face significant difficulties in pronunciation.

The Level of Difficulty Students Experience in Pronunciation

Based on the collected data, various levels of difficulty in pronouncing the consonant sounds zh, ch, sh, d, and t were found among the students. It was identified that three students experienced a very high level of difficulty in pronouncing these sounds. This difficulty may be attributed to several factors, such as a lack of phonetic understanding, habitual mispronunciation, or articulation disorders. Additionally, eleven students were found to still face difficulties, although their level of difficulty was not as severe as the first group. This group requires special attention and more intensive practice to significantly improve their pronunciation skills. Meanwhile, only one student did not experience any difficulty in pronouncing the consonant sounds zh, ch, sh, d, and t. This indicates that 93.33% of the students still need further guidance and practice to enhance their pronunciation abilities.

From these findings, it can be concluded that the focus of instruction should be directed towards improving the pronunciation of the sounds zh, ch, sh, d, and t, especially for students experiencing moderate to high levels of difficulty. Appropriate approaches, such as pronunciation exercises, the use of audio-visual media, and interactive teaching methods, are expected to help students overcome these challenges, enabling them to pronounce these sounds accurately and fluently. The questionnaire results indicate that students experience the greatest difficulty distinguishing between the sounds zh and ch, particularly when others pronounce them. Students find it challenging to differentiate zh from ch due to their similarity and the nearly identical tongue placement, often resulting in confusion between the two consonants.

Additionally, the questionnaire revealed that students have considerable difficulty both pronouncing and distinguishing between zh-ch and t-d. The zh consonant, which requires specific tongue skills and closely resembles the ch consonant, causes most students significant difficulty in pronunciation and listening comprehension. Similarly, for the consonant t, which requires an

aspirated breath, students frequently fail to produce the necessary aspiration, while conversely, they sometimes incorrectly aspirate the consonant d.

Types of Pronunciation Difficulties

The greatest pronunciation difficulties were found in retroflex sounds, particularly in the consonants zh and ch. Based on the analysis of students' pronunciation errors, several specific patterns of difficulty were identified. A total of six students experienced difficulties in distinguishing and producing the consonants *zh* and *ch*, indicating confusion between these two similar retroflex sounds. In addition, six students showed difficulties specifically with the consonant *zh*, suggesting that this sound posed the greatest challenge in terms of articulation. One student experienced difficulty with both *zh* and *sh*, reflecting problems in differentiating closely related retroflex consonants. Another student had difficulty only with the consonant *ch*, while one student experienced difficulty with the consonants *zh* and *t*, indicating confusion between a retroflex consonant and a dental stop. Overall, these findings highlight that the consonant *zh* was the most problematic sound for students, frequently contributing to pronunciation errors either independently or in combination with other consonants.

The following are the most common pronunciation difficulties identified among the students, with a focus on retroflex and aspirated consonant sounds. First, the sounds zh, ch, and sh are commonly reported as difficult for foreign learners of Mandarin, particularly due to their articulatory similarity and specific tongue positioning requirements. These consonants are classified as retroflex sounds, which require the tongue to curl backward toward the roof of the mouth during pronunciation. All students agreed that producing these sounds correctly is challenging because of the deeply curled tongue position. For learners whose first language is Indonesian, this articulatory movement is unfamiliar, making the pronunciation of these sounds especially difficult.

Secondly, among these consonants, the zh sound was the most frequently mispronounced. Students often produced zh as sh or ch, mainly because

DOI: dx.doi.org/10.21274/ls.2025.17.2.349-378

their tongue was not placed far enough back in the mouth. Some students positioned their tongues too shallowly, resulting in unclear articulation. In addition to incorrect tongue placement, several students struggled with maintaining fluency while pronouncing zh, with noticeable hesitation or stuttering occurring during speech production. Further, the ch sound, which is an aspirated retroflex consonant, also posed significant difficulties for students. Many learners were not accustomed to pronouncing sounds or words that contain ch, leading to unclear or inaccurate pronunciation. As a result, students sometimes substituted ch with more familiar consonants such as c or s. This substitution not only affected pronunciation accuracy but also had the potential to change word meanings or reduce intelligibility.

Finally, the consonant pairs zh–ch and d–t were identified as particularly confusing for students. The sounds zh and ch share very similar tongue placements, with the primary difference being that zh requires a slightly more retracted tongue position. Meanwhile, the distinction between d and t lies in aspiration: t is aspirated and produced with a noticeable burst of air, whereas d is non-aspirated. These subtle phonetic differences contributed to students' difficulties in accurately distinguishing and producing the sounds.

Factors Contributing to Difficulties

These pronunciation difficulties are caused by a combination of interlinguistic and extralinguistic factors. Interlinguistic factors include the influence of the Indonesian phonological system, which does not recognize retroflexion and aspiration as meaning-distinguishing features, making it difficult for students to position the tongue and control the airflow accurately. For example, when pronouncing zh or ch, the tongue must be curled toward the roof of the mouth, a position unfamiliar to Indonesian speakers. Some students reported that their tongues felt "too curved," making it difficult to control articulation.

Extralinguistic factors include a lack of intensive practice outside the classroom, low motivation, and limited exposure to authentic Mandarin sounds. Although some students routinely watch music videos or listen to

audio from native speakers, others rarely practice pronunciation outside class hours. The lack of articulatory awareness also causes students to tend to repeat Indonesian pronunciation patterns, resulting in consistent phonological errors.

Without intensive practice, it is difficult to pronounce the consonants zh, ch, sh, d, and t accurately. This lack of involvement directly impacts the low accuracy of pronunciation, especially for sounds that do not exist in the Indonesian language. Listening practice or watching Mandarin-language films has proven helpful, but consistent practice is needed to support progress. The lack of practice may be caused by various factors such as low motivation, insufficient support, limited time, and frequent laziness. Without encouragement from within themselves or external factors, students tend to revert to the articulation patterns of their native language, resulting in pronunciation that does not conform to Mandarin phonology. Therefore, intensive practice is highly necessary both within the school environment and beyond.

Indonesian Language Interference

The difficulties in pronouncing Mandarin consonants among Indonesian-speaking students are primarily caused by differences in articulation, aspiration, and consonant types. Students tend to simplify or substitute unfamiliar sounds with those more familiar in their native language, such as pronouncing the consonant /zh/ as /c/ (Zhàoxiàng becomes Chaoxiang). Consonants like /zh/ and /sh/ require the tongue to be curled backward, a position uncommon in Indonesian, making them difficult to articulate accurately. Moreover, Indonesian does not distinguish aspirated sounds, so consonants like /t/ are pronounced without a strong burst of air. As a result, students often produce /t/ and /ch/ without aspiration, causing them to sound weak or unclear to native Mandarin speakers for example, /t/ being pronounced as /d/ (dong becomes tong). Differences in consonant types also present challenges, as Mandarin includes fricative consonants such as /zh/ and /sh/ and plosive consonants like /d/ and /t/, whereas Indonesian lacks /zh/, rarely uses /sh/ naturally, and has no retroflex consonants. Consequently,

DOI: dx.doi.org/10.21274/ls.2025.17.2.349-378

students tend to equate these sounds or pronounce them inconsistently due to a lack of familiarity.

CONCLUSION

This study analyzed the pronunciation difficulties of the consonants zh, ch, sh, d, and t among ninth-grade students at SMPIT Cheng Hoo and identified the factors causing these difficulties. The results indicate that the retroflex sounds zh, ch, and sh are the most challenging consonants for students to pronounce. Most students substitute them with sounds more familiar in Indonesian, such as pronouncing zh as j, ch as c, and sh as s. Additionally, errors also occur in distinguishing aspirated and non-aspirated consonants, particularly between t and d. For example, the word tā (he/she) is often pronounced as da. These pronunciation errors not only affect communication clarity but also potentially cause shifts in meaning.

The factors causing pronunciation difficulties include: (1) linguistic factors, namely interference from Indonesian, which does not differentiate between retroflex and aspirated sounds; (2) physiological factors, referring to students' limitations in controlling tongue position and speech organs when producing certain sounds; and (3) psychological factors, such as nervousness and lack of confidence, causing students to speak hastily. These three factors interact and reinforce the occurrence of substitution and articulation errors.

The findings underscore the importance of more explicit phonetic instruction in Mandarin language learning at the junior high school level. Teachers need to provide detailed explanations of the differences between retroflex, aspirated, and non-aspirated sounds, as well as train students through articulation exercises, the use of audio-visual media, and repeated corrections. With these strategies, it is hoped that students will improve pronunciation accuracy and build confidence in communicating using Mandarin. Nevertheless, this study has several limitations.

First, the research focused on a limited set of consonants, and therefore did not assess pronunciation difficulties in other sounds or in

aspects such as Mandarin intonation and rhythm. Second, the duration of student observation and practice was limited, meaning the long-term effectiveness of the recommended instructional strategies could not be fully evaluated. Based on these limitations, future research is recommended to investigate pronunciation difficulties in a variety of communicative contexts. Subsequent studies could also explore the influence of learning motivation, independent practice strategies, and the use of audiovisual-based technology on students' pronunciation abilities. In this way, the findings would be more comprehensive and could serve as a foundation for developing more effective teaching methods.

DOI: dx.doi.org/10.21274/ls.2025.17.2.349-378

REFERENCES

Afrina, U., & Celeveresty, T. B. (2020). Analisis Kesalahan Pelafalan Bunyi Konsonan (Z, C, S, Zh, Ch, R) Pada Mahasiswa Program Studi D-4 Bahasa Dan Budaya Tiongkok Angkatan 2017 Universitas Padjajaran. Metahumaniora, 79-89.

Amin, N. F., Garancang, S., & Abunawas, K. (2023). Konsep Umum Populasi Dan Sampel Dalam Penelitian. Jurnal Pilar, 1978-5119.

Anggito, A., & Setiawan, J. (2018). Metodologi Penelitian Kualitatif. Yogyakarta: CV Jejak.

Anggreani, Marsuki, R. R., Amanda, A., & Irawati, R. P. (2021). Analisis Dialek Mandarin Tiongkok dan Mandarin Taiwan dari Film The Ex-File 3 : The Return of The Exes (Tiongkok) dan Film Our Times (Taiwan). Longda Xiaokan: Jurnal of Mandarin Learning and Teaching, 12-17.

Awaru, T. (2024). Kesalahan Pelafalan Bahasa Mandarin Di Kalangan Mahasiswa Program Studi Bahasa Mandarin Dan Kebudayaan Tiongkok Universitas Hasanuddin Pronounciation Errors In Mandarin Among Students Of Mandarin Language And Chinese Culture At Hasanuddin University. Repository UNHAS, 1-60.

Chomsky, N., & Halle , M. (1968). The Sound Pettern Of English. New york, london: Harper & Row.

Ladefoged, P. (2006). A Course In Phonetics. Losh Anggelas: Harcourt Brace Jovanovich College Publishers.

Lado, R. (2007). Re-reading Robert Lado, 1957, Linguistics across Cultures. Applied linguistics for language teachers. International Journal of Applied, 159-276.

Mania, S. (2008). Observasi Sebagai Alat Evaluasi Dalam Dunia Pendidikan dan Pengajaran. Lentera Pendidikan, 220-233.

Milles, M. B., & Huberman, A. M. (1994). Qualitative Data Analysis: An Expanded Sourcebook". London: SAGE Publication.

Mutmainnah, R., Azizah, L., & Junaeny, A. (2023). Junior High School Students' Ability of Chinese Initial and Finals Pronounciation in Speaking. Sinoligua: Journal of Chinese Studies, 39-50.

Najmah, Adeliani, N., Sucirahayu, C. A., & Zanjabila, A. R. (2023). Analisis Tematik Pada Penelitian Kualitatif. Jakarta Selatan: Salemba Medika.

Nurhusnah, N., Rahmanita, R., & Sakinah, Z. (2024). Difficulties of Javanese Ethnic Students Pronouncing Mandarin Vocabolary in The Chinese Languange and Culture Program a Universitas Sebelas Maret, Surakarta. Journal of Maobi Linguistics, Education, Literatur, and Culture, 176-182.

Sugiarti, T., Kusumaningtyas, D. P., Phanata, S., Rudiansyah, & Lianisyah, U. Y. (2022). Analisis Pelafalan Fonologi Konsonan Bahasa Mandarin Z,C,Zh,Ch Mahasiswa Semester 2 Program Studi D-3 Bahasa Mandarin Sekolah Vokasi Universitas Sebelas Maret. Jurnal Cakrawala Mandarin, 1-13.

ting, y. (2011). Bunyi konsonan bahasa indonesia dan bahasa mandarin: analisis kontastif. jurnal bahasa dan sastra, 126-136.

Weirench. (1953). Languages in Contact: Findings and Problems. The Hague (Den Haag), Netherlands: Mouton & Co., The Hague.

Yuliani, W., & Siliwangi, I. (2018). Metode Penelitian Deskriptif Kualitatif Dalam Perspektifbimbingan Dan Konseling. Quanta, 1-9.

Zaid, L. N., & Putra, M. (2024). Dasar-dasar fonetik dalam pembelajaran bahasa. jurnal yayasan haiah nusratul islam, 88.

DOI: dx.doi.org/10.21274/ls.2025.17.2.349-378