Syllabic Distribution and Developmental Patterns of Mandarin Glides in Preschool Children

I-Ping Wan, Xiang Li, Yu Ching Tsai

Proceedings of the 39th Pacific Asia Conference on
Language, Information and Computation (PACLIC 39)

Emmanuele Chersoni, Jong-Bok Kim (eds.)

2025

Syllabic Distribution and Developmental Patterns of Mandarin Glides in Preschool Children

I-Ping Wan*, Xiang Li, Yu Ching Tsai
Graduate Institute of Linguistics, National Chengchi University, Taipei, Taiwan
ipwan@mail2.nccu.tw, 113555010@g.nccu.edu.tw,
113555004@g.nccu.edu.tw

Abstract

This study investigates the developmental patterns of three Mandarin glides, [j], [w], and [ɥ], in preschool children by examining their distribution across different syllabic positions. Data were collected from 45 typically developing children between the ages of 7 months and 6 years using a cross-sectional picture-naming task. The speech produced by the children was transcribed using IPA and compared with adult usage patterns extracted from a 202-hour spoken corpus of Taiwan Mandarin. Statistical analyses, including chi-square tests and linear mixed-effects models, showed that [j] and [w] were acquired earlier and produced more accurately than [ɥ]. Glide accuracy was significantly higher in onset positions and lower in coda positions and consonant-glide clusters. These findings suggest that articulatory complexity and positional effects both influence the course of glide acquisition in Mandarin. By including a comprehensive range of syllabic contexts, this study highlights the importance of structural distribution in early phonological development and provides empirical support for phonotactic modeling in child language acquisition. This study combines corpus-based and experimental evidence to address glide acquisition in a tonal language, focusing on phonological modeling, speech development, and child language research.

1 Introduction

Previous studies have shown that the acquisition of sounds in early childhood is accomplished through stages of development. The trajectory of phonological development in early ages is said to be influenced by language universal factors (e.g., Jakobson, 1968), markedness (e.g., Edwards, 1974), and the articulation ease (e.g., Locke, 1972). Certain patterns in phonological development appear to be cross-linguistically universal, while others are shaped by the specific linguistic input children receive, and still others are influenced by articulatory effort.

Numerous studies have examined the acquisition order of English glide sounds. Dodd et al. (2003) found that [w] tends to be acquired earlier than [j], consistent with earlier findings by Wellman (1931), Poole (1934), Templin (1957), and Sander (1972). To account for the variation in age of acquisition across these studies, Dodd et al. (2003) suggested that both age and gender significantly influence speech development, contributing to differences in observed acquisition timelines. Despite variability in specific ages reported, these studies consistently agree on the developmental trajectory of English glides, with [w] acquired before [j].

Based on the relationship between surface glides and their underlying vowel counterparts in Mandarin phonology, the developmental sequence of these glides can be inferred from the acquisition patterns of the high vowels. Prior research has consistently shown that the high front vowel /i/ emerges first in child speech, followed by the high back vowel /u/, and finally the high front rounded vowel /y/ (Jeng, 1979; Wu & Xu, 1979; Su, 1985; J. Hsu, 1987; Cao, 2003; Shi & Wen, 2007; H. Hsu, 2016). This developmental progression, from simpler to more complex articulatory gestures, aligns with Locke's (1972) theory of articulatory ease. Accordingly, if glides reflect the acquisition order of their vowel counterparts, it can be hypothesized that palatal rounded [j], derived from high front vowel /i/, is acquired earlier than labial-velar rounded [w] and palatal rounded [ɥ], which are derived from /u/ and /y/, respectively.

Previous studies on Mandarin phonological development have consistently shown that lip rounding presents greater articulatory difficulty for young children compared to lip spreading or neutral postures. For example, Peng & Chen (2020) and Lou (2020) reported that rounded vowels and glides ([w], [ɥ]) are typically acquired later than [j], reflecting higher motor control demands for rounding. Similarly, Zhang (2016) observed frequent substitution and deletion of [w] and [ɥ] in Mandarin-Taiwanese Min children, indicating articulatory instability. From a phonological perspective, Wiedenhof (2015) and Fu (2023) noted that [ɥ] has a restricted distribution and is often merged with [j], suggesting that the Mandarin phonological system itself limits the robustness of [ɥ]. Even non-native speakers of Mandarin, Thai preschoolers, learn Mandarin was found the similar patterns (Wan et al, 2024a).

There are twelve different syllable structure combinations in Mandarin (V, CV, GV, VG, VN, CVG, CVN, GVG, GVN, CGV, CGVG, CGVN). G represents the three glides [w, ɥ, j] in Mandarin. Except for the nasal sound /n/, all other consonants are strictly patterned to be in syllable initial position; glides, on the other hand, are more flexible (Chao, 1968; C.C. Cheng, 1973; Duanmu, 2007). They can appear in syllable initial (e.g., jaw), medial (e.g., ɕjaw), and final positions (e.g., xaj); however, it is not legitimate for [ɥ] to occur in syllable-final position (Duanmu, 2007; Norman, 1988). Table 1 below shows the twelve syllable structure types in Mandarin:

Syllable structure	Phonetic Transcription	Gloss
V	[i 55]	clothes
CV	[pi 21]	pencil
GV	[ja 35]	teeth
VG	[wɔ 21]	I, me
VN	[i 51]	hard
CVG	[naj 21]	milk
CVN	[san 21]	umbrella
GVG	[jow 21]	to have
GVN	[wan 21]	bowl
CGV	[xwa 55]	flower
CGVG	[njow 35]	cow
CGVN	[twan 21]	short

Table 1: The twelve syllable structure types in Mandarin

Although Mandarin allows some flexibility, it imposes strict phonotactic constraints on consonant-glide onset sequences. As noted by Wan (1999), glides show position-specific behavior in terms of both manner and place of articulation. The glide [w] co-occurs with bilabials, dentals, velars, and retroflexes; [j] occurs with bilabials, dentals, and palatals; and [ɥ] appears primarily with palatals and the nasal [n]. Table 2 shows the possible distribution of CG sequences by Wan (2003) shows the legal CG segment structures in Mandarin (see Appendix I). For the co-occurrence of vowels and glides, it is undoubtedly that the three high vowels /i/, /y/, and /u/ cannot combine with the three glides, due to the homorganicity tautosyllabic constraints (Steriade, 1988). For instance, [ij], [yɥ], and [uw] are not allowed. According to Wan and Jaeger (2003), non-high vowels can be paired with the three glides freely in the pattern of VG sequence combinations. For example, [aj], [ɥɛ], [ow], [fej], and [xaw].

Mandarin includes three glides: the labiovelar [w], the palatal [j], and the labiopalatal [ɥ]. These are generally considered surface realizations of the high vowels /u/, /i/, and /y/, respectively, when they occur adjacent to non-high vowels. None of these glides are part of the underlying phonemic inventory of Mandarin. Instead, they emerge through phonological processes that affect high vowels in specific syllabic environments. This interpretation is supported by several studies (Lin, 1989; Wu, 1994; Duanmu, 2002; Wan, 1999; Wan 2002; Fu 2023) and aligns with the view that glides in Mandarin function as allophonic variants of high vowels rather than as independent semivowels or consonants.

Wan (2002, 2003, 2006) presents a series of studies on glide behavior in Mandarin speech errors, offering compelling evidence that glides function more like nuclear elements than true consonants. First, glide errors revealed that [j], [w], and [ɥ] show positional variation and often align with the nucleus, particularly when preceded by labial, dental, or retroflex consonants. In contrast, glides form onset clusters when following palatal or velar consonants. Wan and Jaeger (2003) further observed that glides frequently substitute for their high-vowel counterparts and may trigger vowel-vowel substitutions across syllables. These interactions suggest that glides share more representational features with vowels than with consonants or other glides. Importantly, glide errors such as substitution, deletion, and insertion often occur independently, indicating their segmental status. In a follow-up study, Wan (2006) analyzed postnuclear glides and coda nasals in native speaker speech errors. Glides showed significantly fewer interactional errors than nasals (79 vs. 158), and interacted more with vowels than with consonants. Although glides occasionally co-occurred with nasals, the pattern suggested that glides are less stable as codas and more tightly integrated with the nucleus. Collectively, these findings support the analysis of glides as prosodically and phonologically affiliated with the nucleus, rather than as independent onset or coda consonants.

In English, only vowel monophthongs can form diphthongs or triphthongs, while glides such as [j] and [w] are typically analyzed as consonants (Ladefoged and Maddieson, 1996). In contrast, Mandarin diphthongs and triphthongs are formed through the combination of a glide, which is considered the surface form of the high vowels [i], [u], or [y], and one or more non-high vowels. This contrast highlights a fundamental difference in the phonological treatment of glides in the two languages. In Mandarin, the behavior and distribution of glides correspond more closely to the properties of vowels than to those of consonants. Examples of diphthongs in Mandarin and token frequencies are shown as below.

IPA	Freq.	IPA	Freq.	IPA	Freq.
ɕja	18031	swan	1727	ʂwən	337
ɕjɛn	14461	tʂwan	1498	lwan	324
tɕja	11453	tʂ^{h}wan	1496	tɕ^{h}ɥən	307
mjɛn	10427	tɕ^{h}ja	1337	tɕɥən	287
pjɛn	9907	kwa	1226	k^{h}wan	281
tɕ^{h}jɛn	8286	tʂwa	1171	ʐwan	228
tɕjɛn	8224	k^{h}wa	1045	twən	228
tjɛn	8019	ɕɥən	972	swən	216
njɛn	7820	xwa	724	tswən	214
t^{h}jɛn	6324	tʂwən	688	nja	118
lja	5760	xwən	682	tɕ^{h}jo	115
kwan	5689	ɕjo	641	kwən	55
tɕ^{h}ɥen	3705	tʂ^{h}wa	619	nwan	54
xwan	2990	t^{h}wən	576	t^{h}wən	49
ljɛn	2766	tɕɥɛn	466	ʐwən	39
ɕɥen	2452	tʂ^{h}wən	453	ʂwan	27
twan	2096	t^{h}wan	378	tswan	20
lwən	1874	ʂwa	361	ts^{h}wan	3
p^{h}jɛn	1806	k^{h}wən	348	tɕjo	1

Table 3: Sample of CGVN in IPA and token frequencies (Wan et al., 2024b)

Since glide development in Mandarin has typically been examined within the broader context of vowel development, the specific behavior and acquisition of the three glides [w], [j], and [ɥ] in different syllabic positions remain insufficiently investigated in the literature. This study addresses the following research questions, each accompanied by a predicted outcome:

1. What is the developmental order of the three Mandarin glides [j], [w], and [ɥ] in preschool children between 7 months and 6 years of age? Based on previous studies, we predict that [j] and [w] will emerge earlier, while [ɥ], due to its articulatory complexity, will be acquired later

2. How do syllabic positions (onset, post-consonantal, and coda) affect the accuracy and emergence of Mandarin glide production in child speech? We expect that glides will be produced more accurately in onset position than in post-consonantal or coda positions, reflecting structural facilitation in early speech.

3. To what extent do articulatory complexity and positional effects influence the acquisition patterns of Mandarin glides? It is predicted that the interaction of higher articulatory demands (as in [ɥ]) and marked syllabic positions will result in greater variability and delayed mastery compared to simpler contexts.

Glide acquisition in child speech has often been examined with limited positional scope, typically focusing on onset production. However, Mandarin glides present a unique opportunity to examine positional asymmetries, given their attested occurrences in onset, coda, and complex cluster contexts. Mandarin prohibits vowel sequences with a single syllable, high vowels are systematically realized as glides when adjacent to other vowels. Therefore, this study aims to address these gaps by analyzing the production accuracy of [j], [w], and [ɥ] across syllabic positions in child speech, and comparing these patterns with large-scale adult corpus data. Drawing on markedness theory and positional prominence, we ask how phonotactic constraints shape developmental trajectories, and how corpus-driven benchmarks can refine our understanding of normative phonological development.

2 Methodology

The data were drawn from a spoken corpus of Taiwan Mandarin children (N = 45; mean age = 3.8 years, SD = 1.4; 23 boys, 22 girls) collected in the Phonetics and Psycholinguistics Laboratory at National Chengchi University, Taipei, Taiwan [1].

[1] In this lab, the spoken corpus contains multi-tier linguistic annotations and encompasses diverse spoken data such as speech interactions among multiple speakers, conversations between speech therapists and Mandarin-speaking aphasic patients, language acquisition patterns in typically developing children aged 7 months to 6 years, and speech samples from children with language disorders aged 3 to 6 years. More recently, the corpus has been expanded to include not only

Participants, aged 7 months to 6 years, completed picture-naming tasks designed to elicit three target glides [j], [w], and [ɥ] in both monosyllabic and disyllabic words across different syllabic positions, as shown in Table 4

Word	IPA	Word	IPA
羊	[ja35]	月亮	[ɥɛ51 lja51]
碗	[wan21]	貝殼	[pej51 kʰɤ35]
雲	[ɥən35]	帽子	[maw51 tsɨ]
海	[xaj21]	樹葉	[ʂu51 jɛ51]
猴	[xow35]	青蛙	[tɕʰi55 wa55]
掃把	[saw51 pa21]	醫院	[i55 ɥɛn51]
牙齒	[ja35 tʂɨ21]	牛奶	[njow35 naj21]
襪子	[wa51 tsɨ]	蛋糕	[tan51 kaw55]

Table 4: The test examples in the experiment

Participant information is shown in Table 5 (see Appendix II). All data were audio-recorded and transcribed by two well-trained master's students specializing in phonetics, who had prior experience transcribing over 24 languages before undertaking this task, with interrater 91%, and 9% of discrepancies were resolved through acoustic validation in Praat. To enhance reliability, Praat was used to analyze the recordings, ensuring consistency in transcriptions and validating any discrepancies. When confusions still remained, the data were discarded. Annotation was conducted semi-automatically using a Hybrid-DNN-HMM framework. For developmental analysis, children were grouped in six-month intervals (2;0–6;0), enabling observation of age-related changes in phonological development, with vocabulary production serving as an index of phonological maturity.

All pictures were presented three times in a randomized order. As preschool children may not always produce the exact target word, approximations were accepted as correct as long as the target glides were realized. For example, [tɕʰi55 wa55] *'frog'* produced as [wa55 wa55] or [maw51 ts] *'hat'* as [maw51 maw51] were considered correct because the target [w] occurred. Similarly, [njow35 naj21] *'milk'* produced as [nej55 nej 55] and [xaj21] *'sea'* produced as [wej21] *'water'* were also counted as correct, since the target items involved a glide [j] in coda position. However, if the target [w] did not appear in the production, such as [tan51 kaw55] *'cake'* produced as [tan51 tan51], the response was coded as incorrect. A total of 84 test items involving true consonants and glides across different syllabic positions were examined. In this cross-sectional design, if the participant was unable to produce the target word correctly, the response was taken to indicate that the child at that age had not yet mastered the sound. The procedures followed those outlined in Wan et al. (2024b). Table 6, as seen in Appendix III, shows the correctness and error types by the 45 participants.

3 Findings and Analysis

Data were collected through a picture-naming task involving typically developing Mandarin-speaking children in Taiwan. The task was designed to elicit productions of the glides [j], [w], and [ɥ] across various syllabic positions. Of the 7,290 expected tokens, 6,555 usable responses were obtained, with the shortfall due to omissions during the task. Among these, 6,292 productions were judged accurate, yielding an overall accuracy rate of 95.99%. Three main error types were identified in the remaining data: substitution (1.16%), where the intended glide was replaced; addition (0.26%), where an extra glide was inserted; and deletion (2.59%), where the glide was omitted. Productions containing multiple error types were excluded from analysis to ensure categorical clarity.

The glide [w] appeared most frequently in the confusion matrix, with 3,110 total instances, of which 3,102 were produced accurately. Substitution errors were rare and included realizations such as [j] (n = 1), [p] (n = 7), [pʰ] (n = 1), [v] (n = 1), and [] (n = 8). Deletion of [w] was observed in 106 cases. These results indicate a high level of production accuracy for [w], although some substitution and deletion errors did occur. Table 7 shows the error type distribution according to the three glide sounds [j], [w], [ɥ] in Mandarin.

Glide / Error Type	Substitution	Deletion	Addition	Total Error
j (34.1%)	7	62	14	83
w (51.8%)	18	106	3	127
ɥ (14.1%)	33	2	0	35
Total	**58**	**170**	**17**	**245**

Table 7: Error type distribution according to the three glide sounds

This table reveals that among the three glides, [w] accounts for the highest proportion of errors (51.8%), with deletion being the most frequent type (106 instances). This suggests that [w] may be particularly vulnerable to deletion during speech pro-

data from adult learners of Mandarin with L1 backgrounds in Indonesian, Vietnamese, and Thai, but also speech data from native speakers of these three languages.

duction. The glide [j] also shows a predominance of deletion errors, but with a relatively higher number of additions, indicating its potential instability or overgeneralization in certain phonological contexts. In contrast, [ɥ] is primarily affected by substitution errors, with almost no deletions or additions, implying that it may be more prone to misidentification or confusion with other sounds. Overall, deletion emerges as the most common error type, highlighting the articulatory challenges glides may pose in speech development. The observed and expected numbers of the three glide error types below show the extent to which the observed distribution deviates from the expected frequencies under the assumption of independence, forming the basis for the chi-square calculation.

Glide	Error Type	Observed	Expected
[j]	Substitution	7	19.65
[j]	Deletion	62	59.59
[j]	Addition	14	5.76
[w]	Substitution	18	30.07
[w]	Deletion	106	88.12
[w]	Addition	3	8.81
[ɥ]	Substitution	33	8.29
[ɥ]	Deletion	2	24.29
[ɥ]	Addition	0	2.43

Table 8: Expected frequencies and observed frequencies regarding the error type distribution

By examining the relationship between the three glides ([j], [w], and [ɥ]) and error types (substitution, addition, and deletion) through the chi-square test of independence, results reveal that a significant correlation can be observed ($\chi^2(4) = 128.87, p < .001$). In the test, the degrees of freedom (df = 4) were calculated, based on the categories in both variables, the three glides and the syllable position in this case. The test then compares the expected and observed frequencies to see if the variables are related. Overall, the distribution of errors is dependent on the glide type, suggesting that the errors in children's output is systematically related to the syllable position of the three glides.

$$df = (rows - 1) \times (columns - 1) \quad (1)$$

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad (2)$$

A closer observation on the three syllable positions regarding the three glides, [ɥ] is the only sound prohibited from occurring at the syllable-final position. For [w], its occurrence in the syllable-final-position was the highest. As for [j], it appears the most in the consonant-glide cluster. The total count of production according to the syllable position is 6361. Again, [ɥ] showed minimal distribution and variation, with only two syllable positions allowed in Mandarin phonotactics. Combining Table 7 and 8, the pattern of Mandarin glide syllable distribution (Table 9) can be observed: [ɥ] had the lowest overall frequency, but it showed the highest error rate ($35/292 \approx 11.99\%$); meanwhile, even though [w] and [j] had a higher occurrence compared to [ɥ], the error rate was 3.99% and 2.33% separately. According to each glide sound, the significant error types vary as well. [ɥ] exhibited the most instances of substitution; [w] had the highest deletion count; [j] experienced a high addition rate. The results from Table 7 and 8 suggest that both the occurrence frequency and syllable distribution influence the Mandarin glide production in normal children.

IPA	Syllable Position	Number of Production	Examples
[w] 3110 (48.89%)	CG	768	[xwa55]
	syllable-initial	258	[wan21]
	syllable-final	2084	[tɕʰjow35]
[j] 2959 (46.52%)	CG	1711	[ljɛn21]
	syllable-initial	373	[jaŋ35]
	syllable-final	875	[tswej21]
[ɥ] 292 (4.59%)	CG	83	[ɕɥoŋ35]
	syllable-initial	209	[ɥən35]
Total count		6361	

Table 9: Mandarin glides syllable distribution and examples

In Table 9, [[ɕɥoŋ35] "bear" became [tjoŋ35]. under the cases of substitution, [xwa55] "flower" became [fa55]. The glide [w] was deleted and the consonant [x] was replaced by [f]. The disappearance and replacement of the glide [w] and the consonant [x] showed a closer relationship between the onset and prenuclear glide. Substitutions like [tjow35] "ball" became [tj35], [o] and [w] were substituted with [ɔ]. The relationship between the postnuclear glide [w] and the nucleus [o] is bounded. They were substituted together by mid vowel [ɔ]. The substitution example satisfies the observation she found: the tautosyllabic prenuclear glide, vowel, and postnuclear glide are treated as a constituent, backed up by the ability to be substituted with a single vowel. Consequently, postnuclear glides should be considered part of the nucleus rather than the coda.

Building on their position, the status of glides can be further suggested. The classification of glides cannot be uniform in all positions. Their status is syllable-sensitive, shaped by the structural interaction with the adjacent sounds within one syllable. The following table below shows the error types and syllable distribution of the three glides in Mandarin.

IPA	Syllable Position	Number of error	Examples
[w] (124, 54.39%)	CG	28	[tswej]→[tsej] [ja ʂwa]→[xa ʂja]
	syllable-initial	9	[wan]→[pan]
	syllable-final	87	[tɕjawʌɨ]→[tɕja] [njaw]→[njaŋ]
[j] (69, 30.26%)	CG	18	[ɕjaŋ tɕjaw]→[xaŋ tjaw]
	syllable-initial	5	[ja ʂwa]→[xa ʂwa] [ja ʂwa]→[a twa]
	syllable-final	46	[pej kʰɤ]→[pe] [xaj]→[xaə]
[ɥ] (35, 15.35%)	CG	7	[ɕɥoŋ]→[ɕjoŋ] [ɕɥoŋ]→[ʂoŋ]
	syllable-initial	28	[i ɥɛn]→[i jan]
Total count	228		

Table 10: Error types and syllable distribution

This table analyzes the distribution of pronunciation errors for the glides [w], [j], and [ɥ] across different syllable positions. The glide [w] shows the highest error rate, accounting for 54.39% of all errors, with the majority occurring in syllable-final position. The glide [j] ranks second with 30.26% of errors, mainly found in syllable-final and CG structures. The glide [ɥ] has the fewest errors (15.35%), appearing only in CG and syllable-initial positions, and not at all in syllable-final positions. These findings highlight the varying stability of glides in different phonological environments, with [w] being particularly vulnerable in syllable-final contexts. Table 11 (see Appendix IV) shows correctness and all the error types, along with the percentage out of all produced data.

The confusion matrix (see Appendix VII) helps further identify specific error patterns of the three glides, [w], [j], and [ɥ], revealing which glides are most frequently misarticulated and the phonemes with which these glides most often interact.

4 Discussion and Conclusion

Regarding the developmental order of the three Mandarin glides [j], [w], and [ɥ] in preschool children aged between 7 months and 6 years, the consistently high accuracy rates for both [w] and [j] indicate that these glides are more stable in child speech than [ɥ], which emerged later and was less accurate. Substitution errors were observed only from [ɥ] to [w] or [j], but not in the opposite direction, suggesting an asymmetrical developmental pattern. The frequency of correct productions followed the order [w] > [j] > [ɥ], aligning with Wan's (2003) distributional ranking of legal CG sequences and with adult spoken corpus data reported in Wan et al. (2024b). The correlation between children's and adults' production rankings further suggests that early glide acquisition is strongly shaped by frequency distributions in the ambient language.

In terms of syllabic positions (onset, post-consonantal, and coda), these factors significantly affect the accuracy and emergence of Mandarin glide production in child speech. Chi-square analysis confirmed a significant relationship between glide type and error type ($\chi^2(4) = 128.87$, $p < .001$), indicating that errors were not random but systematically linked to syllable position. Positional analysis showed that [w] frequently occurred in syllable-final position, [j] predominated in consonant–glide clusters, and [ɥ] was limited to initial and medial positions. Error patterns also varied across contexts: [ɥ] was most prone to substitution, [w] and [j] to deletion, and [j] exhibited a higher addition rate among the three sounds. These findings reflect the positional distribution of glides and demonstrate that syllable structure strongly conditions glide accuracy and emergence.

Finally, to some extent, articulatory complexity and positional effects influence the acquisition patterns of Mandarin glides. The findings confirm that articulatory demands and positional restrictions jointly shape glide acquisition. [w] and [j] were acquired earlier and used more consistently due to their positional flexibility and lower articulatory complexity. In contrast, [ɥ], which requires lip rounding in addition to palatal articulation, was more unstable, frequently substituted, or deleted. This supports prior research showing that lip rounding is generally more difficult than lip spreading for Mandarin-speaking children (Peng & Chen, 2020; Lou, 2020; Zhang, 2016; Wan et al., 2024a). The structural behavior of glides further corroborates earlier phonological models: in syllable-initial position, glides function as onsets or secondary articulations (Duanmu, 1990, 2002), while in syllable-final position, postnuclear glides pattern with the nucleus (Lin, 1989). Overall, the interaction of articulatory complexity and syllable structure accounts for both the variability and the delayed mastery observed in [ɥ] relative to [j] and [w].

In addition, our findings are consistent with prior research showing that lip rounding is generally more difficult for Mandarin-speaking children than lip spreading. Rounded glides such as [ɥ] and [w] tend to emerge later, are often substituted by [j], or deleted altogether, reflecting both articulatory constraints on lip rounding and phonological restrictions in early child Mandarin (Peng & Chen, 2020; Lou, 2020; Zhang, 2016; Wan et al. 2024a).

These findings suggest that the acquisition pattern in all language systems is shaped by a combination of a universal developmental pattern, the language-specific sound inventory, and the articulatory constraints. To better understand the systematic trajectory of sound acquisition, it is essential to investigate the status of individual sound types within a particular language.

Acknowledgments

We sincerely appreciate the valuable and constructive comments from the two anonymous reviewers and the editor, which have significantly improved this manuscript. All remaining errors in the analysis and interpretation are solely our own. This research was supported in part by the National Science and Technology Council in Taiwan (NSC 100-2410-H-004-187-) to the first author.

References

Bao, Z. (1990). Fanqie languages and reduplication. *Linguistic Inquiry*, 21(3), 317-350.

Bao, Z. (1996). The syllable in Chinese. *Journal of Chinese Linguistics*, 24, 312-354.

Baxter, W. H. (1992). *A handbook of old Chinese phonology*. Berlin, New York: De Gruyter Mouton. DOI:10.1515/9783110857085.

Cao, J. X. (2003). One case study of early phonological development in Mandarin-speaking children. In *Proceedings of the 6th National Symposium on Modern Phonetics*. Tianjin Normal University.

Chao, Y. R. (1934). The non-uniqueness of phonemic solutions of phonetic systems. Bulletin of Institute of History and Philology. *Academia Sinica*, 4(4), 363-397. DOI:10.6355/BIHPAS.193401.0363

Chao, Y. R. (1948). The voiced velar fricative as an initial in Mandarin. *Le Maitre Phonetique*, 63, 2-3.

Chao, Y. R. (1968). *A grammar of spoken Chinese*. Berkeley: University of California Press.

Cheng, C. C. (1973). *A Synchronic Phonology of Mandarin Chinese.* Berlin, New York: De Gruyter Mouton. DOI:10.1515/9783110866407

Cheng, R. L. (1966). Mandarin phonological structure. *Journal of Linguistics*, 2(2), 135–158. DOI:10.1017/S0022226700001444

Chiang, W. Y. (1992). *The prosodic morphology and phonology of affixation in Taiwanese and other Chinese languages.* [Doctoral dissertation, University of Delaware]. ProQuest Dissertations Publishing.

Chung, R. F. (1989). *Aspects of Kepa phonology.* [Doctoral dissertation, University of Illinois]. ProQuest Dissertations Publishing.

Chung, H., Kong, E. J., Edwards, J., Weismer, G., Fourakis, M., & Hwang, Y. (2012). Cross-linguistic studies of children's and adults' vowel spaces. *The Journal of the Acoustical Society of America*, 131(1), 442–454. https://doi.org/10.1121/1.3651823

Davis, S., & Hammond, M. (1995). On the status of onglides in American English. *Phonology*, 12(2), 159-182. DOI:10.1017/S0952675700002463

Davis, B. L., & MacNeilage, P. F. (1990). Acquisition of correct vowel production: a quantitative case study. *Journal of Speech and Hearing Research*, 33(1), 16–27. https://doi.org/10.1044/jshr.3301.16

Duanmu, S. (1990). *A Formal Study of Syllable, Tone, Stress, and Domain in Chinese Languages.* [Doctoral dissertation, MIT]. ProQuest Dissertations Publishing.

Duanmu, S. (2002). *The Phonology of Standard Chinese.* Oxford University Press.

Duanmu, S. (2007). *The phonology of standard Chinese (2nd ed.).* New York: Oxford University Press Inc.

Dodd, B., Holm, A., Hua, Z., & Crosbie, S. (2003). Phonological development: a normative study of British English-speaking children. *Clinical Linguistics & Phonetics*, 17(8), 617–643. https://doi.org/10.1080/0269920031000111348

Donegan, P. (2002). Phonological processes and phonetic rules. *Future Challenges for Natural Linguistics*, 57-81.

Edwards, M. L. (1974). Perception and production in child phonology: The testing of four hypotheses. *Journal of Child Language*, 1(2), 205-219.

Fu, B. (2023). Uncovering Mandarin Speaker Knowledge with Language Game Experiments (Doctoral dissertation, Massachusetts Institute of Technology).

Gussman, E. (2007). *The Phonology of Polish.* Oxford University Press.

Hartman, L. M. (1944). The segmental phonemes of the Peiping dialect. *Language*, 20, 28-42.

Hockett, C. F. (1947). Peiping Phonology. *Journal of the American Oriental Society*, 67, 211-222.

Hockett, C. F. (1950). Peiping morphophonemics. *Language*, 26, 63-85.

Howie, J. M. (1976). Acoustical studies of Mandarin vowels and tones (Vol. 18). Cambridge University Press.

Hsu, H. Y. (2016). *Phonological development and disorder in Taiwan Mandarin: The status of glides* (Master's thesis). National Chengchi University.

Hsu, J. (1987). A *study of the various stages of development and acquisition of Mandarin Chinese by children in Taiwan milieu* [Unpublished Master's dissertation]. Fu Jen Catholic University.

Ingram, D. (1989). *First language acquisition: Method, description and explanation.* Cambridge University Press.

Jakobson, R. (1968). *Child language: aphasia and phonological universals (No. 72).* Walter de Gruyter.

Jeng, H.-H. (1979). The acquisition of Chinese phonology in relation to Jakobson's law of irreversible solidarity. In *Proceedings of the 9th International Congress of Phonetic Sciences: Vol. 2* (pp. 155–161). Copenhagen, Denmark: University of Copenhagen.

Karlgren, B. (1954). *Compendium of Phonetics in Ancient and Archaic Chinese.* SMC Pub. Inc.

Kim, C. W., & Kim. H-Y. (1991). The character of Korean glides. *Studies in the Linguistic Sciences,* 21(2), 113-125.

Ladefoged, P., & Maddieson, I. (1996) *The sounds of the world's languages.* Oxford: Blackwell Publishers.

Li, F., & Munson, B. (2016). The development of voiceless sibilant fricatives in Putonghua-speaking children. *Journal of Speech, Language, and Hearing Research,* 59(4), 699-712.

Lin, Y. H. (1989). *Autosegmental Treatment of Segmental Process in Chinese Phonology.* [Doctoral dissertation, University of Texas]. ProQuest Dissertations Publishing.

Locke, J. L. (1972). Ease of articulation. *Journal of Speech and Hearing Research,* 15(1), 194-200. https://doi.org/10.1044/jshr.1501.194

Lou, S. (2020). Early Phonological development in Mandarin: An analysis of prosodic structures, segments and tones from babbling through the single-word period (Doctoral dissertation, University of York).

McLeod, S. (2009). Speech sound acquisition. In J. E. Bernthal, N. W. Bankson & P. Flipsen Jnr (Eds.), *Articulation and phonological disorders: Speech sound disorders in children* (6th ed., pp. 63-120 + 385-405). Boston, MA: Pearson Education.

McLeod, S., & Crowe, K. (2018). Children's consonant acquisition in 27 languages: A cross-linguistic review. *American Journal* *of Speech-Language Pathology*, 27(4), 1546-1571.

Miao, X., & Zhu, M. (1992). Language development in Chinese children. *Advances in Psychology*, 90, 237-276. https://doi.org/10.1016/S0166-4115(08)61894-4

Moser, D. (1991). Slips of the tongue and pen in Chinese. Department of Oriental Studies, University of Pennsylvania.

Norman, J. (1988). *Chinese*. Cambridge University Press.

Peng, G., & Chen, F. (2020). Speech development in mandarin-speaking children. In Speech perception, production and acquisition: Multidisciplinary approaches in Chinese languages (pp. 219-242). Singapore: Springer Singapore.

Poole, I. (1934). Genetic development of articulation of consonant sounds in speech. *The Elementary English Review*, 159-161.

Priester, G. H., Post, W. J., & Goorhuis-Brouwer, S. M. (2011). Phonetic and phonemic acquisition: Normative data in English and Dutch speech sound development. *International Journal of Pediatric Otorhinolaryngology,* 75(4), 592-596. https://doi.org/10.1016/j.ijporl.2011.01.027

Sander, E. K. (1972). When are speech sounds learned?. *Journal of Speech and Hearing Disorders,* 37(1), 55-63.

Scullen, M. E. (1993). *The Prosodic Morphology of French.* [Doctoral dissertation, Indiana University]. ProQuest Dissertations Publishing.

Shen, J. (1993) Slips of the tongue and the syllable structure of Mandarin Chinese. InS-C. Yau (ed.) Essays on the Chinese Language by Contemporary Chinese scholars. Paris: Centre de Recherches Linguistiques sur l'Asie Orientale-Ecole des Hautes Etudes en Sciences Sociales. 139-162.

Shi, F., & Wen, B. Y. (2007). Vowel development in Mandarin-speaking children. *Zhongguo Yuwen, 2007(5)*, 444–454.

Steriade, D. (1988). Review of CV Phonology: A Generative Theory of the Syllable, by G. N. Clements & S. J. Keyser. *Language*, 64(1), 118–129. DOI:10.2307/414790

Su, A-T. (1985). *The acquisition of Mandarin phonology by Taiwanese children.* [Master's dissertation]. Fu Jen Catholic University.

Templin, Mildred C. (1957). *Certain language skills in children: Their development and interrelationships.* Minneapolis: University of Minnesota Press.

Wan, I. P. (1999). *Mandarin phonology: Evidence from speech errors.* [Doctoral dissertation, State University of New York]. ProQuest Dissertations Publishing.

Wan, I. P. (2002). The Status of Prenuclear Glides in Mandarin Syllables: Evidence from Psycholinguistics and Experimental Acoustics. 聲韻論叢, (11), 141-162. https://doi.org/10.29753/CP.200210.0008

Wan, I. P. (2003). *Alignments of prenuclear glides in Mandarin. Taipei*: Crane Publishing.

Wan, I. P. (2006). A psycholinguistic study of postnuclear glides and coda nasals in Mandarin. *Journal of Language and Linguistics*, 5(2), 158-176.

Wan, I. P., Allassonnière-Tang, M., & Yu, P. (2024a). Early Segmental Production in Thai Preschool Children Learning Mandarin. *International Journal of Asian Language Processing*, 34(02), 2450005.

Wan, I. P., Chang, C. W., Lee, C., & Yu, P. (2024b). Probability Distributions of Sounds and Phonotactics in Taiwan Mandarin Syllables. In *Proceedings of the 38th Pacific Asia Conference on Language, Information and Computation* (pp. 1157-1165).

Wan, I. P., & Jaeger, J. J. (2003). The phonological representation of Taiwan Mandarin vowels: A psycholinguistic study. *Journal of East Asian Linguistics, 12*(3), 205-257. DOI:10.1023/A:1023666819363

Wellman, B. L., Case, I. M., Mengert, I. G., & Bradbury, D. E. (1931). *Speech sounds of* *young children.* University of Iowa Studies: Child Welfare.

Wiedenhof, J. (2015). Appendix A The International Phonetic Alphabet. In A Grammar of Mandarin (pp. 409-409). John Benjamins Publishing Company.

Wu, Y. (1994). M*andarin Segmental Phonology.* [Doctoral dissertation, University of Toronto]. ProQuest Dissertations Publishing.

Wu, T. M., Xu, Z. Y. (1979). A preliminary analysis of language development in children during the first three years. *Acta Psychologia Sinica, 11*(2), 153–165.

Zhang, J. Q. (2016). Nonsibilant Fricative Acquisition by Bilingual Guoyu-Taiwanese Southern Min Children (Master's thesis, The Ohio State University).

Zhu, H. (2002). Phonological Development in Specific Contexts: Studies of Chinese-speaking. Blue Ridge Summit: Multilingual Matters. https://doi.org/10.21832/9781853595899

Zhu, H. & Dodd, B. (2000). The phonological acquisition of Putonghua (modern standard Chinese). Journal of Child Language, 27(1), 3-42.

Appendix

I. Possible distribution of CG sequences by Wan (2003)

Table 2 shows the possible distribution of CG sequences by Wan (2003) shows the legal CG segment structures in Mandarin: asterisk * indicates a CG sequence is not possible.

	Bilabial	**Labial**	**Dental**	**Retroflex**	**Palatal**	**Velar**
Unaspirated plosive	pj, pw		tj, tw			*kj, kw
Aspirated plosive	p^hj, p^hw		t^hj, t^hw			*k^hj, k^hw
Fricative		*fj, *fw	*sj, sw	*ʂj, ʂw ʐj, ʐw	ɕj, ɕɥ, *ɕw	*xj, xw
Unaspirated affricate			ts^hj, ts^hw	*tʂj, tʂw	tɕj, tɕɥ, *tɕw	
Aspirated affricate				*$tʂ^h$j, $tʂ^h$w	$tɕ^h$j, $tɕ^h$ɥ, *$tɕ^h$w	
Nasal	mj, mw		nj, nɥ, nw			*j, *w

Table 2: The possible distribution of CG sequences by Wan (2003)

II. Participant information

Table 5 shows all 45 participant information.

III. Correctness and error distribution

Table 6 shows the correctness and error distribution types by the 45 participants.

ID	Gender	Age
1	M	0；7；9
2	F	1；5；22
3	F	1；8；2
4	M	1；8；28
5	F	1；9；19
6	M	2；4；0
7	F	2；4；5
8	M	2；4；21
9	M	2；5；1
10	F	2；5；18
11	M	2；6；8
12	M	2；9；13
13	F	2；9；24
14	M	2；10；6
15	F	2；10；8
16	F	3；0；9
17	M	3；1；24
18	M	3；4；26
19	M	3；5；28
20	M	3；7；17
21	F	3；9；22
22	F	3；10；11
23	M	3；10；23
24	M	3；11；20
25	F	4；1；25
26	M	4；2；26
27	M	4；5；4
28	M	4；5；15
29	F	4；6；23
30	F	4；6；29
31	F	4；7；27
32	F	4；8；0
33	F	4；8；7
34	M	4；8；13
35	F	5；1；4
36	M	5；2；0
37	F	5；2；23
38	M	5；2；25
39	M	5；3；27
40	F	5；8；4
41	F	5；8；4
42	M	5；9；11
43	M	5；9；21
44	F	5；9；27
45	F	5；10；26
SD=1.38；Average age=3.8		

Table 5: Participant information

IV. Error Types

Table 11 shows correctness and all the error types, along with the percentage out of all produced data

V. Error examples

The table of the proportion of error examples for each of the three glides below (Table 12) helps provide a clear data showing which phoneme each of the three glides is most frequently substituted with during the replacement process.

VI. Confusion Matrix

The confusion matrix below helps further identify specific error patterns of the three glides, [w], [j], and [ɥ], revealing which glides are most frequently misarticulated and the phonemes with which these glides most often interact.

Participant	Correctness	Substitution	Deletion	Addition	Item counts
1	137	0	5	2	144
2	47	1	3	1	52
3	128	3	18	2	151
4	72	6	15	0	93
5	93	3	21	0	117
6	151	0	2	0	153
7	112	2	6	0	120
8	97	18	21	0	136
9	116	0	7	0	123
10	160	1	0	0	161
11	146	13	7	0	166
12	136	1	8	0	145
13	156	2	1	1	160
14	146	0	10	2	158
15	113	0	1	2	116
16	147	6	1	1	155
17	122	0	11	0	133
18	155	1	2	0	158
19	104	0	6	2	112
20	151	0	0	0	151
21	147	1	9	0	157
22	150	0	5	0	155
23	115	1	2	0	118
24	162	0	0	0	162
25	152	0	2	0	154
26	160	0	0	0	160
27	155	1	0	2	158
28	161	0	1	0	162
29	158	0	0	0	158
30	162	0	0	0	162
31	138	6	0	2	146
32	160	0	0	0	160
33	158	0	0	0	158
34	108	2	0	0	110
35	162	0	0	0	162
36	162	0	0	0	162
37	158	0	0	0	158
38	150	0	2	0	152
39	162	0	0	0	162
40	157	0	1	0	158
41	147	5	1	0	153
42	157	1	0	0	158
43	156	0	0	0	156
44	144	2	2	0	148
45	162	0	0	0	162
Total	**6292**	**76**	**170**	**17**	**6555**

Table 6: Correctness and error distribution by participant

	Correctness	%	Substitution	%	Deletion	%	Addition	%	Sum(counts)
j	2898	97.22%	7	0.23%	62	2.08%	14	0.47%	2981
w	3102	96.07%	18	0.56%	106	3.28%	3	0.09%	3229
ɥ	292	89.30%	33	10.09%	2	0.61%	0	0.00%	327
Total	6292	96.25%	58	0.89%	170	2.60%	17	0.26%	6537

Table 11: Error types

IPA	Correctness	%	Error Types	Error counts	%	Sum
j	2898	98%	j→w	1	0.067%	2967
			j→ɥ	0	0.000%	
			j→x	2	0.067%	
			j→i	2	0.067%	
			j→ə	1	0.034%	
			j→y	1	0.034%	
			j→#	62	2.090%	
w	3102	96%	w→j	1	0.031%	3226
			w→ɥ	0	0.000%	
			w→p	7	0.217%	
			w→p^h	1	0.031%	
			w→v	1	0.031%	
			w→ŋ	8	0.248%	
			w→#	106	3.286%	
ɥ	292	89%	ɥ→j	21	6.422%	327
			ɥ→w	1	0.306%	
			ɥ→n	1	0.306%	
			ɥ→y	10	3.058%	
			ɥ→#	2	0.612%	

Table 12: The proportion of error examples for each of the three glides

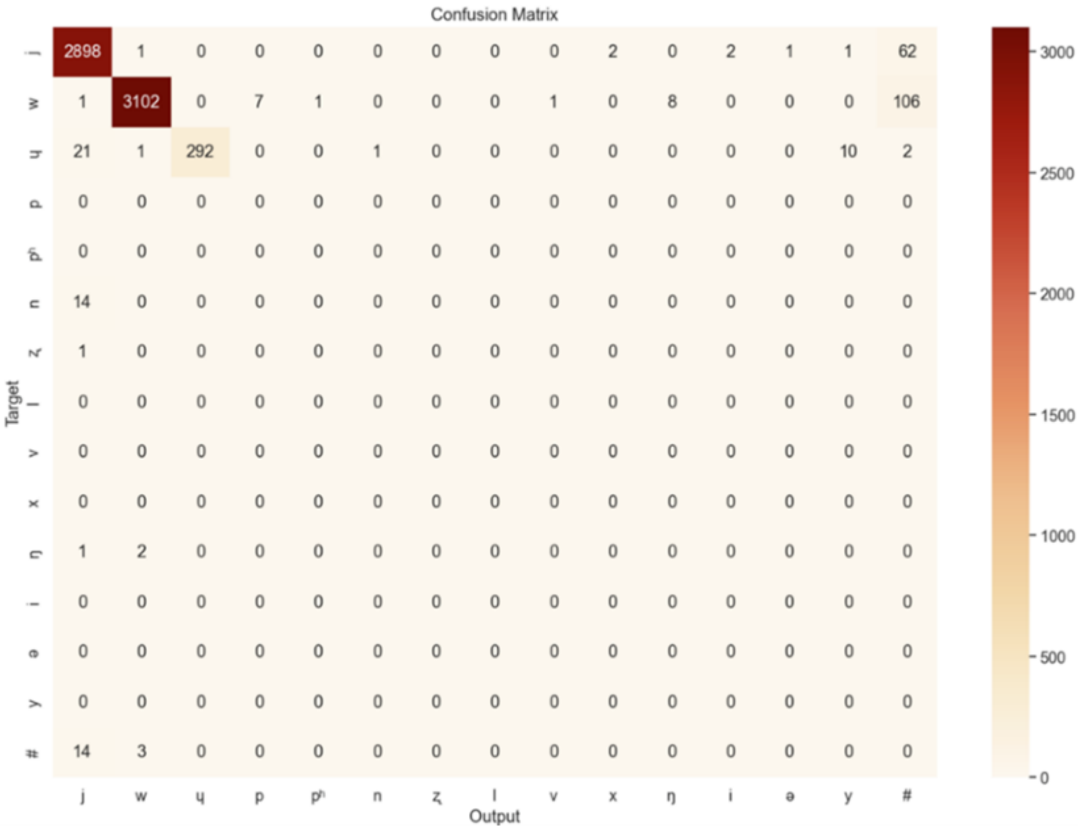

Figure 1: Confusion matrix (indicates that a phone is dropped, a deletion phenomenon)