Vowel-glide Distinction in High Vocoid Diphthong Structures in Squliq Atayal: An ultrasound and acoustic study

Huang Bo-Xuan

Graduate Institue of Linguistics

National Taiwan University

This paper provides the first ultrasound investigation in the vowel-glide distinction in Squliq Atayal, and shows, in combination with acoustic data, that both Squliq Atayal /ui/ and /iu/ structures are GV structures, with the second segments being the nuclei. Crucially, the tongue contours of /u/ in /iu/ and the vowel /u/ do not show statistical difference, and both the intensity and duration of the second segments are higher/longer than the first, indicating a rising in sonority. This paper hopes to pave way for future ultrasound studies of glide–vowel distinction, and to provide novel insights into the syllable structures of Squliq Atayal.

keywords: Squliq Atayal, glide, vowel, ultrasound, tongue configuration

1. Introduction

The similarity between and the defining criteria for high vowels and glides have long been the center of attention in both the fields of phonetics and phonology (e.g. Rosenthall, 1994; Delattre, 1966; Chitoran, 2002; Levi, 2011; Jaggers, 2018). In the field of phonology, some works propose using either syllabic structure (Clements & Keyser, 1983; Kaye & Lowenstamm, 1984) or phonological features of [±syllabic] and [±sonorant] (Chomsky & Halle, 1968) for the distinction between glides and high vowels. While in many languages such as Latin, such a distinction is sufficient (cf. 1, adopted from Levi, 2011, where glide/vowel can be predicted according to the syllabic structure), in many other languages, syllabification only solves half the puzzle, as there exist situations where both readings are possible, and Squliq Atayal is such a language. In Squliq Atayal, syllable structures are minimally CV and maximally CGVC, and no consonant or vowel clusters are allowed (de Carvalho, 2015). This immediately raises the question as to how a vocoid cluster structure should be analyzed, as one must surface as the nucleus and the other a glide (cf. (2)).

(1) a. V→G/_V#:

/iecur/ [je.kur] 'liver'
/ouis/ [o.wis] 'I come'

b. V→V/elsewhere:

/mulier/ [mu.li.er] 'woman'
/dies/ [di.es] 'day'

(2)		/βui/	[βuj] or [βwi]	'shoot.IMPER.'
		/ləliu/	[ləliw] or [ləlju]	PN
	cf.	/otoβai/	[otoβaj]	'motocycle'

While this issue is rather easily resolved when one vocoid is /a/, as it has no glide counterpart and the other high vocoid must surface as glide, we are faced with the conundrum of how to analyze the /iu/ and /ui/ structures, where both vocoids can surface as the glide, and a phonological consensus has been lacking regarding which should be the case and how we can determine it. In Hsu (2004a; 2004b)'s studies on Taiwan Southern Min and Taiwan Hakka, it is found that /i/ remains as the nucleus when two syllables with /u/ and /i/ respectively contract, which she, along with evidences from rhyming patterns, takes to be proof that /i/ is the nucleus in /ui/ and /iu/ syllables. H.-F. Yang (personal communication, April 16, 2021), however, disapproves by arguing that properties of monophthongs and diphthongs are not on par with each other, and such a cross-comparison may not be convincing.

On the other hand, phonetic accounts regarding the measurement of glides and vowels, and the results of it, are not consistent, either. While Huang (2014) proposes that in Squliq Atayal, the structure of /iu/ is a sequence of a glide followed by vowel, reflecting a rising in sonority, Hsu (2004a; 2004b) find that in Taiwan Southern Min and Taiwan Hakka, it is /i/ that carries higher F2 amplitude than /u/, which is taken to be indication of higher sonority. Apart from sonority, nucleus and glides are also suggested to contrast each other in duration. Comparing the durations of nucleus and glides in three other Formosan languages, Wu (2002) argues that glides in these languages behave like a true consonant, which links directly to the syllable node, differing in terms of duration from a true vowel, which is linked to a mora.

Since a consensus is lacking from phonological point of view, and phonetic, especially articulatory data are sparse, this paper draws evidence from ultrasound tongue configuration in production of the structures in question, with support from acoustic evidences as well. To the best of the authors' knowledge, it is the first ultrasound study examining the glide-vowel distinction in Squliq Atayal, and wishes to provide novel insight from an articulatory point of view.

2. Method

2.1 Participant

One naïve female native speaker of Squliq Atayal (Mekarang) in her 70s participated in the data collection, with no previous history of speech disorder upon self-report. The participant was born in Mekarang, and moved to Nahuy in her 20s. She is therefore fluent in both Squliq and Ci'uli Atayals.

2.2 Materials and Procedures

To observe the behaviors of the vowels and the targets, the vowels /u/, /i/ and the diphthongs /ui/[1], /iu/ are taken as the stimuli. To make comparison with the glides, /aw/, /wa/, and phonologically disambiguous /ju/ (e.g. /tɕjux/ Progressive marker) and /wi/ (eg. /rawin/ 'friend') are also taken. All the targets are in the final syllables of the words, which are the stressed syllables in Squliq Atayal, and said in the carrier phrase *Kmal saku' ke'* ___. 'I say the speech of ___.' Please refer to Appendix 2.2 for a full list.

[1] It should be noted that in Squliq Atayal orthography, word-final /ui/ can theoretically be written as *uy* or *wi* and /iu/ as *yu* or *iw*. While for /iu/, we did not find any word ending with /iw/, for /ui/,words ending in *uy* or *wi* are plenty. To disperse the concern that orthographical diference might indicate phonological difference, a *post hoc* examination was done, which showed no difference in terms of the tongue configuration of the /ui/ written as *uy* or *wi*.

2.3 Apparatus

The ultrasound images were recorded with a portable ultrasound machine (CGM OPUS 5100), through a transvaginal electronic curved array probe (CLA 651). The transducer was fixed perpendicular to the tongue floor on the midsagittal plane by a headset. A USB 3.0 powered capture card (ExtremeCap U3) was used to record the ultrasound data, with the frame rate set at 37 fps. The data were saved as .mp4 files. The scanning rate was 6.5 MHz and the reception frequency ranges from 4 to 9 MHz at a scan depth of 8.8 cm. The audio data were collected with a microphone (Audio–Technica carcoid AT2035) plus a portable audio interface (USBPre 2).

2.4 Statistical analysis

The segments were first labeled in Praat (Boersma & Weenink, 2009) manually. The tongue contours of the segments were then traced with Matlab's GetContours, and fitted through generalized additive mixed models (GAMMs; Wieling, 2018).

As for the acoustic materials, the F1 and F2 were extracted with Python's Parselmouth (Jadoul et al., 2003) automatically for each segments, with 101 points each. Duration of each segment was calculated with the F1 or F2 as the indicator. As the transition from the first segment to the second is continuous and dynamic in the segments investigated in this study, F2 was used as indicator for the /iu/, /ui/, /ju/, /wi/, /ju:/, and /wi:/ segments, and F1 for the /wa/, /aw/, /ja/, and /aj/ segments, as F2 is conventionally regarded as a function of anteriority–posteriority of the vowels, which contrasts /i/ and /u/ from each other, while F1 varies according to the height of the vowels, which contrasts /j/ and /w/ from /a/. The mean F2s of the single vowels /i/ and /u/, and mean F1s of /i/, /u/, and /a/ were used as the thresholds for the /ui/–/iu/ segments and glide segments respectively. The duration of a target in the segment was calculated by first adding the difference between the F1 or F2 value at each of the 101 points and the targets' mean F1 or F2 values, which were then divided by the difference of the F1 or F2 values between the two items in the targets (e.g. /i/ and /u/ in /iu/, or /w/ and /a/ in /wa/). This was taken as the percentage of how similar each of the 101 points was to the target, if by any chance this value exceeded 100%, it was taken as 100%. An example of how the duration of /i/ in /iu/ and /ui/ was calculated is shown in (3), where D_i is the duration of /i/, $F2_n$ is the F2 at the point, $\overline{F2\_i}$ and $\overline{F2\_u}$ are the respective mean F2 values of monophthong /i/ and /u/, and D is the duration of the token.

(3)

$$D\_i = \sum_{n=1}^{101} \frac{F2_n - \overline{F2\_i}}{\overline{F2\_u} - \overline{F2\_i}} \times D$$

This method results in having the sum of the respective durations of the two segments longer than the original duration, that is, for example, the durations of /u/ in /ui/ and /i/ in /ui/ added up would become longer than /ui/'s duration, by about 2-5 ms. The final determined duration is therefore set as the midpoint of the overlap of the two segments' durations.

3. Results

3.1 Ultrasound results

Figure 1 shows the tongue contours of the /i/ group and the /u/ group. It is immediately seen that the /i/ group segments have an overall identical configuration, while the contours of the /u/ group segments tell more about the potential status of the segments. When comparing /u/ in /ui/, /u/ in /iu/, and monophthong /u/, we see that /u/ in /ui/ is the most anterior amongst the three, followed by /u/ in /iu/ and finally /u/, as shown in Figure **??**. This suggests that /u/ in /iu/ is closer to the vowel than /u/ in /ui/ is. Crucially, /u/ in /iu/ shows no difference with the monophthong /u/, as in Figure **??**, and /u/ in /ui/ is also rather identical to the glide /w/ in /wi/, as in Figure **??**, with no difference at their major regions of constriction, indicating that the /u/ in /iu/ is more vowel-like, and the /u/ in /iu/ has an overall glide-like configuration.

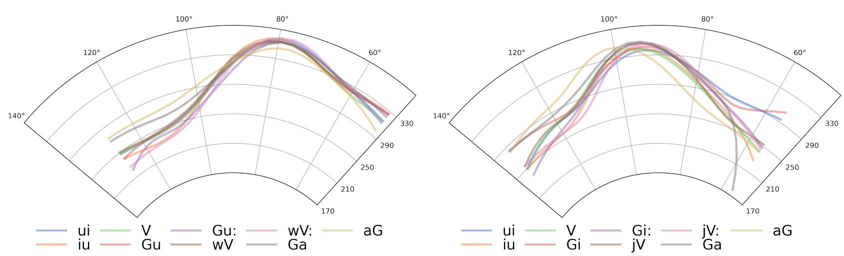

Figure 1: GAMMs results: (a) /i/ group; (b) /u/ group.

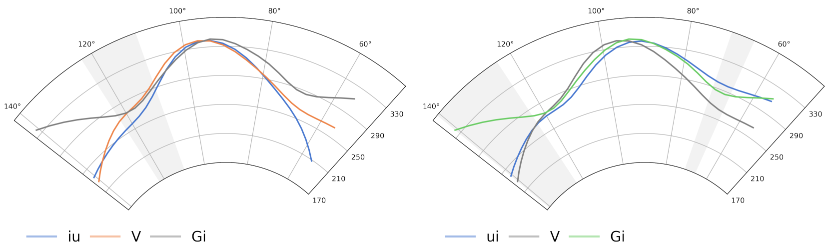

Figure 2: GAMMs comparison: /u/ in /iu/ vs. /u/ and /u/ in /iu/ vs. /w/.

3.2 Duration

The duration results show that when the second segment is not the glide, or is ambiguous in terms of its vowel-glide status, it is longer than the preceding vocoid. If we take Wu (2002)'s account that duration implies difference between glides and vowels, our result seems to suggest that when a vocoid combination is ambiguous for both GV or VG interpretations, GV is preferred. That is to say, in both /ui/ and /iu/ structures, it is the latter vocoid that is longer in duration, and thus more likely to be the nucleus in Squliq Atayal.

3.3 Intensity

Another often considered criterium when evaluating sonority is intensity (Parker, 2002). Generally, the higher the intensity of a segment is, the higher it is in the sonority hierarchy. Figure **??** shows results similar to the duration findings, where the latter parts of the investigated components, if not obligatorily glides, are generally higher in intensity. As can be seen in Figure 4, if we

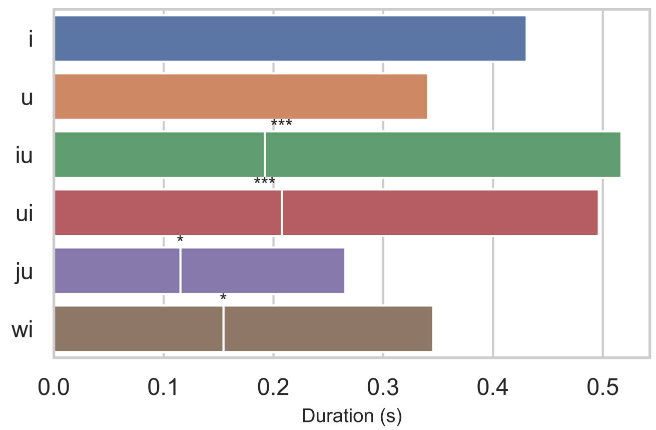

Figure 3: Duration of each segment. The boundaries between the first and second segments are indicated by the white lines in the bars. The asterisks indicate where the durations within the groups are significant (t-test). It is seen that whichever group it is, the latter segment, if not a glide, is longer than the former.

put duration and intensity together, it comes out clear that in both /ui/ and /iu/ sequences, it is the second halves of the combinations that have higher intensity than the preceding vocoids, suggesting a larger emphasis on the latter vocoid than the former.

3.4 Sectional summary

This section investigates the segments from ultrasound and acoustic data, with two main findings. /i/ group elements in general have rather identical tongue shapes, even between glides and the vowel /i/, while the /u/ group segments are more variable, and the /wi/'s /w/ and /ui/'s /u/ show no difference with each other at their major constriction regions, both being more in the front than the vowel and the /u/ in /iu/. Acoustic findings including duration and intensity also

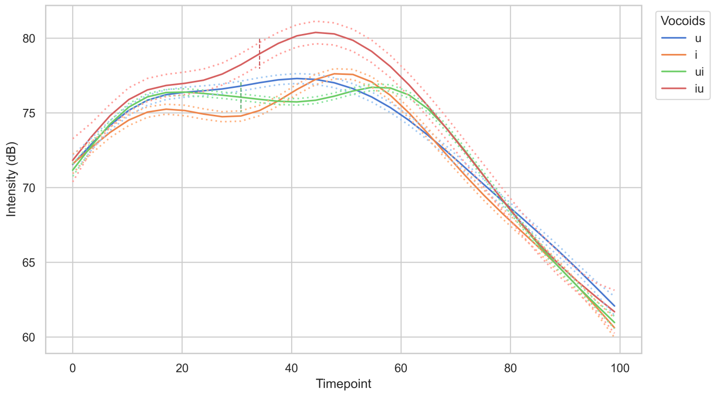

Figure 4: GAMMs results: Intensity of /ui/ and /iu/. Dashed lines indicate segment boundary. It is seen that the latter halves of the diphthongs have higher intensity.

support the view that in both /ui/ and /iu/ sequences, it is the second vocoids that are more nucleus-like.

4. Discussion

In this section, we will discuss the implications of the results, and propose explanations from articulatory and cross-linguistic perspectives.

In the analyses in § 3.4, the conclusion is drawn that in Squliq Atayal, the word-final /ui/ and /iu/, which can theoretically serve as both GV or VG, as both forms would abide by Squliq Atayal's CGVX structure, should be considered as a GV, rather than VG, sequence, since acoustically, both of their intensity high points fall on the second halves, which also have longer duration, and articulatorily, the former component of /iu/ is closer to a glide, the latter a vowel. The two sequences are thus both open syllables.

However, it remains to be discussed why in the /i/ group, we do not see such a difference, that is, why /ui/'s and /iu/'s /i/ are not different from each other, if the former is a glide and the latter a vowel. It should be noted that it is not only the /i/ in /ui/ and /iu/ that are not different, but all the /i/ components we investigate do not show much systematic difference. It is highly probable that it is due to /i/'s coarticulatory resistance. In the literature, /i/ is often reported to have higher coarticulatory resistance, which stops /i/ from coarticulation with adjacent sounds. Stone & Vatikiotis-Bateson (1995) argues that articulators that contribute to vocal constrictions resist coarticulatory effects due to need of precision of articulatory control. Reasonably, /i/ would have higher articulatory resistance than /u/ due to its higher constriction. Alone the same line, Faytak et al. (2020) proposes that alveolar sounds, especially /i/, have high articulatory resistance, and may spread their articulatory demands to surrounding sounds. In their study, /in∼iŋ/ are found to be harder to distinguish for both Shanghainess speakers and Beijing Mandarin speakers, while /aN/ structures are easily differentiated, which the authors attribute to /i/'s tendency to maintain its articulation and its invariable acoustic properties. This matches our ultrasound findings where /i/ group data are rather consistent across different segments, while /u/ segments are much less so. Such an articulatory resistance of /i/ components allows them to maintain their tongue contours regardless of their syllabic status. It is thus likely that acoustic cues are more easily accessible when differentiating glide /j/ from vowel /i/ than tongue contour imaging. This, however, is not to say production does not play a role. It is imaginable that parameters such as tongue muscle activation would be different when producing these two sounds. Electromyographic data may thus be of great use in this study.

Having explained the question of the production of /i/ components. We shall now turn to the bigger question of why does Squliq Atayal choose sequential order over segmental phonological difference, that is, why does not Squliq Atayal bestow /u/ and /i/ with consistent syllabic status respect a certain sequential pattern of syllabic structure. This question leads to another question: if it is sequential order that is preferred, why is it GV instead of VG?

Both questions can be answered if we consider cross-linguistic evidences. In most languages, open syllables are preferred over closed ones. In Benson (1998), it is found that Vietnamese speakers, while having closed syllables in their native languages, when speaking English tend to omit the codas in closed syllables, while they rarely append coda to open syllables. The same is also observed in Squliq Atayal, where speakers have the tendency to drop word-final glottal stops.

As for the first question, it should be noted that Squliq Atayal has a relatively small vowel inventory, with only five phonemic vowels /a/, /i/, /u/, /e/, and /o/. In fact, Proto-Atayal only had the first three, with /e/ and /o/ deriving from /ai/ and /au/, respectively (Li, 1980). Even in the present time, dialectal differences exist where some dialects retain the original diphthongs, and some have undergone monophthongization (eg. /seħuj/ vs. /sajħuj/ 'taro', and /qwo/ vs. /qwaw/ 'alchol'). Languages with simpler sound inventories are reported to allow for larger vowel variance, and presumably emphasize the distinction of consonants over that of vowels. Manuel & Krakow (1984) compares the vowel-to-vowel coarticulation in Swahili, English, and Shona, and contends that motor systems are constrained by the requirement of the maintenance of phonemic distinctiveness. Languages with more phonemic vowels would thus have smaller vowel variance/vowel-to-vowel coarticuation than languages with looser vowel spaces. This larger vocalic variance may lead Squliq Atayal speakers to focus more on segmental or suprasegmental components other than vowels, such as the syllabic structures.

5. Conclusion

This study examines the vowel-glide distinction in Squliq Atayal high vocoid diphthong /iu/ and /ui/ structures, and finds that in both structures, the following segments are more nucleus-like, judging from ultrasound and acoustic data. This study is the first to examine such issue with ultrasound imaging, and hopes to pave way for future studies of the articulation of glide-vowel distinction.

References

Benson, B. (1998). Universal preference for the open syllable as an independent process in Interlanguage Phonology. *Language Learning*, *38*(2), 221–235. doi: 10.1111/j.1467-1770.1988.tb00410.x

Boersma, P., & Weenink, D. (2009). *Praat: doing phonetics by computer (version 6.1.31).* Retrieved from `http://www.praat.org`

Chitoran, I. (2002). A perception-production study of Romanian diphthongs and glide-vowel sequences. *Journal of the International Phonetic Association*, *32*, 203–222.

Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper & Row.

Clements, G., & Keyser, S. J. (1983). CV Phonology: A Generative Theory of the Syllable..

de Carvalho, F. O. (2015). Syllable structure and hiatus resolution in Squliq Atayal. *Language and Linguistics*, *16*(3), 347–368.

Delattre, P. (1966). Voyelles Diphtonguées et Voyelles Pures. In *Studies in French and Comparative Phonetics* (pp. 95–106). Berlin: De Gruyter Mouton. doi: 10.1515/9783112416105-012

Faytak, M., Liu, S., & Sundra, M. (2020). Nasal coda neutralization in Shanghai Mandarin: Articulatory and perceptual evidence. *Laboratory Phonology*, *32*(2), 113–121.

Hsu, H.-C. (2004a). Compositional structure of /iu/ and /ui/ in Taiwanese Southern Min revisited. *Language and Linguistics*, *5*(4), 1003–1018.

Hsu, H.-C. (2004b). On the structure of /iu/ and /ui/ in Sixian Hakka. *Journal of Cinese Linguistics*, *32*(1), 129–156.

Huang. (2014). Phonological patterning of prevocalic glides in Squliq Atayal. *Language and Linguistics*, *15*(6), 801–824.

Jadoul, Y., Thompson, B., & de Boer, B. (2003). Introducing Parselmouth: A Python interface to Praat. *Journal of Phonetics*, *71*, 1–15. doi: 10.1016/j.wocn.2018.07.001

Jaggers, Z. S. (2018). Evidence and characterization of a glide-vowel distinction in american english. *Laboratory Phonology*, *9*(1), 1–27.

Kaye, J. D., & Lowenstamm, J. (1984). De la syllabicité. In F. Dell, D. Hirst, & J.-R. Vergnaud (Eds.), *Forme sonore du langage: Structure des représentations en phonologie* (pp. 123–159). Paris: Hermann.

Levi, S. (2011). Glides. In M. van Oostendorp, C. J. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell Companion to Phonology* (pp. 341–366). Mass. & Oxford: Wiley-Blackwell. doi: 10.1002/9781444335262.wbctp0015

Li, P. J.-K. (1980). The phonological rules of Atayal dialects. *Bulletin of the Institute of History and Philology Academia Sinica*, *51*(2), 349–405.

Manuel, S., & Krakow, R. (1984). Universal and language particular aspects of vowel-to-vowel coarticulation. *Haskins Laboratories Status Report on Speech Research*, *SR-77/78*, 69–78.

Parker, S. G. (2002). *Quantifying the sonority hierarchy* (Unpublished doctoral dissertation). University of Massachusetts Amherst.

Rosenthall, S. (1994). *Vowel/glide Alternation in a Theory of Constraint Interaction* (Unpublished master's thesis). McGill University.

Stone, M., & Vatikiotis-Bateson, E. (1995). Trade-offs in tongue, jaw, and palate contributions to speech production. *Journal of Phonetics*, *23*(1-2), 81–100.

Wieling, M. (2018). Analyzing dynamic phonetic data using generalized additive mixed modeling: a tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics*, *70*, 89–116.

Wu, C.-H. (2002). *A Study of Glides in Formosan Languages: Acoustic Evidence for a Constraint-based Approach* (Unpublished master's thesis). National Chengchi University.

Appendix

A. Words used in data collection

/i/	qani 'this'	lozi 'again'	zngi 'to forget'	sinsi 'teacher'	uzi 'also'
/u/	simu '2SG.NOM'	mamu '2SG.GEN'	qu 'NOM'	makuʔ '1SG.GEN'	
/ui/ (*uy*)	seħui'taro'	məʔui 'tired'	məlikui 'male'	βeħui 'wind'	kəmui 'ring'
/ui/ (*wi*)	βərui 'write.IMPER.'	βui 'shoot.IMPER.'	tsɨɣətui 'naggy'	məqlui 'to drift'	
/iu/	ləliu PN				
/ju/	kuju PN	tɕiux 'PRO'	nɕiux 'PRO'	ləpjuŋ 'relative'	ɕjup 'trash'
/wi/	rawain 'friend'	Qawil PN	kwiʔ 'little insect'	trwiʔ 'mulberry tree'	pələquiʔ 'Taiwan cotton rose'
/ju:/	toju: 'soy sauce'	kəɕju:'to borrow'			
/wi:/	sɨβwi: 'to wrap'				
/wa/ or /aw/	qwaw 'alchohol'	wal 'already'	lwa INTERJ.	aβaw 'leaf'	sakaw 'bed'
/aj/ or /ja/	tɕikaj 'a bit'	jajaʔ 'grandma'	βalay 'really'	kja 'to have'	pajan PN