***Taiwan Journal of Linguistics***
Vol. 15.1, 57-101, 2017
DOI: 10.6519/TJL.2017.15(1).3

# PHONETIC REPRESENTATIONS OF THREE *KIM* WORDS IN TAIWANESE*

Pingyou Lin
*Mingdao University*

**ABSTRACT**

This paper examines the phonetic representations of the three Taiwanese words *kim aʔ* 'golden duck', *kim a* 'gold', and *kim ma* 'place name' under the autosegmental framework (Goldsmith 1976). In accordance with the acoustic, articulatory, and perceptual analyses, the durational differences of the three [m]s across the three words have been shown to be different. However, the moraic theory may be limited to the contrast of the three words (e.g., [kim(μμ)ma(μ)] 'gold' vs. [kim(μμ)][ma(μ)] 'place name') though it may be applicable to only the contrast of [CV(μ)][C̲V(μ)] (Japanese *kita* 'arrived') vs. [CVC̲(μ)(μ)][C̲V(μ)] (*kitta* 'sliced'). As a result, the double-empty slots (xx-slots) and the bidirectional assimilation rule are proposed to explain such disparity. In my proposal, I state that the length differences are attributed to phonetic transition across different prosodic boundaries; the durational contrast of the consonant [m] therefore can explicitly be observed from the surface forms of [kim̲][ˀaʔ] 'golden duck' vs. [kim̲$^{ma}$a] 'gold' vs. [kim̲]$^{mm}$[m̲a] 'place name'.

Key words: Taiwanese, phonetic representation, autosegmental framework, liaison, bidirectional assimilation rule, moraic theory

---

* I would like to say thank you to my father who brought me into the field of Taiwanese phonology and always inspired me with non-mainstream innovative perspectives. I would also like to extend my deepest gratitude to Prof. James Myers for allowing me to audit many of his classes and for making many helpful suggestions for my future research. The generous discussions of Wu Ruiwen, Chen Tsungying, Li Yingshing, Lai Yuda, Su Shioufen and many others on this topic are much appreciated. Finally, I am deeply grateful to the two anonymous reviewers for their thoughtful comments on this study.

## 1. INTRODUCTION

In recent decades, research on affixal phonology in Chinese has drawn much attention (Lin 2004), and its analysis has been largely based on Goldsmith's (1976) Autosegmental Phonology. The research on suffixes in Taiwanese has followed this trend as well. Previous literature generally accepts the emergence of liaison (transitional sound in this paper) between a stem and its suffix in Taiwanese but disagrees on a concrete description of this phenomenon. This paper provides a different view on the transitional sound in the suffixed form *kim a* 'gold' in Taiwanese with acoustic, articulatory, and perceptual evidence demonstrated by native speakers. From a comparison among the analyses of *kim a*, *kim aʔ* 'golden duck' and *kim ma* 'Kim(mng)-Ma(tsoo), place name', the phonetic form of *kim a* with its transitional sound can be specified to shed light on the differences across the three phonetic forms. Due to space limitation, tones and tautosyllabic transitional sounds are excluded from the discussion in the rest of this paper.[1]

One of the goals in autosegmental phonology is to provide a reasonable and straightforward account of phonetic representation. Goldsmith (1976:16) stresses, "Autosegmental phonology is an attempt to supply a more adequate understanding of the phonetic side of the linguistic representation." Phonetic representation mainly reflects how speakers recognize words phonetically, namely the exact sounds articulated and perceived by the speakers. Put differently, the phonetic representation of a word is a sequence of phonetic symbols at the surface level after phonological rules are applied.

In previous literature, three different phonological rules have been proposed to explain transitional sounds in Taiwanese words whose stems are followed by a diminutive suffix -a.[2] First, Yip (1980) proposes a

---

[1] The three lexical items are all composed of *kima*; however, the tone of *kim aʔ*$^{\underline{31}}$ is different from that of the other two *kim a*$^{53}$ and *kim ma*$^{53}$. Such distinction of different tones can easily be perceived by native speakers. As to the tautosyllabic case, a transitional sound occurs within a syllable as well. For example, the syllable *ma* 'horse' has a transitional element emerging between the two segments [m] and [a].

[2] This paper primarily concentrates on the segmental changes but not suprasegmental ones triggered by the -a$^{53}$ suffixation. Tones are thus omitted from all transcriptions.

resyllabification rule, which moves a stem-final consonant from a coda position to the onset position of the following syllable. Take the word *kim a* 'gold' for example, its phonetic representation becomes [ki ma] after undergoing resyllabification.[3] Second, an ambisyllabicity rule is proposed by Kao (2003), Tung (1957,1959), and Wang (1991,1995), and this proposal requires the coda consonant of a syllable to be shared by a succeeding syllable; i.e., both syllables govern the consonant at the same time. With the same example *kim a* 'gold', the ambisyllabicity rule changes its phonetic representation to [kima]. Third, most researchers (Ang 1985; Cheng 1997; Cheng and Cheng Xie 1977; Chiang 1992; Chung 1996; Lin 1989; Wang and Kao 2004; Yang 1991; Zhang 1993; among others) suggest a gemination rule, which fills an empty C-slot of a following syllable via rightward spreading of a preceding coda. Consequently, a common CV syllable structure emerges as the product of the process. The application of this rule changes the phonetic representation of *kim a* 'gold' to [kimma].[4]

All of the above three phonological rules aim to address the emerging transitional sound between two morphemes: Resyllabification and ambisyllabicity capture the transitional sound by restructuring syllable organization, while in the gemination proposal the sound is considered the insertion of a timing slot (skeletal slot). With the syllable theory in Clements and Keyser (1983) and the X-tier model in Levin (1985),[5] the three rules can be represented based on autosegmental phonology as follows (see also Lin 2012b:57):

---

[3] A space in square brackets (phonetic transcription) indicates pauses; that is, a silence appears on a syllable boundary. For further discussions on liaison of *kim a* 'gold', please refer to Lin (2006, 2012a, 2012b).

[4] Regarding other related data, syllable-final stops are generally considered as having undergone both the intervocalic voicing rule and the gemination rule, such as /ap a/→[abba] or [apba] 'box', /ts$^h$at a/→[ts$^h$alla] or [ts$^h$atla] or [ts$^h$atda] 'thief', and /tik a/→[tigga] or [tikga] 'bamboo' (please also refer to the discussion of transitional sounds of syllable-final stops in Lin 2012a). Codas that end with high vowels are considered applicable to the gemination rule as well, while non-high vowels are inapplicable to the rule; e.g., /li a/→[liya] 'plum' and /lu a/→[luwa] 'brush' vs. /e a/→[e_a] 'shoes', /o a/→[o_a] 'oyster', /ɔ a/→[ɔ_a] 'taro', and /aʔ a/→[a_a] 'duckling' (glottal stop deletion).

[5] X can either be a C or a V depending on the context. Regardless of tenseness or laxness of vowels and contextual differences in the duration of consonants, every segment

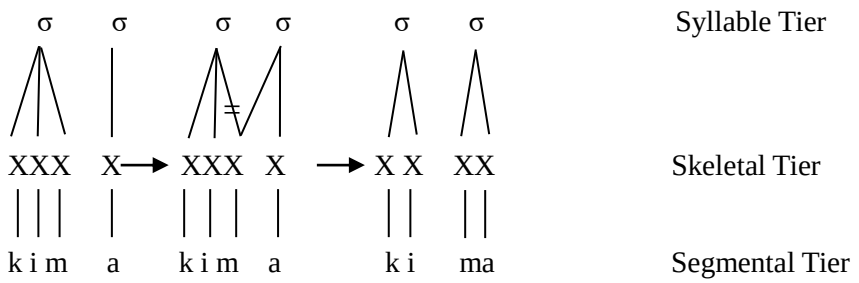



Figure 1. Resyllabification rule

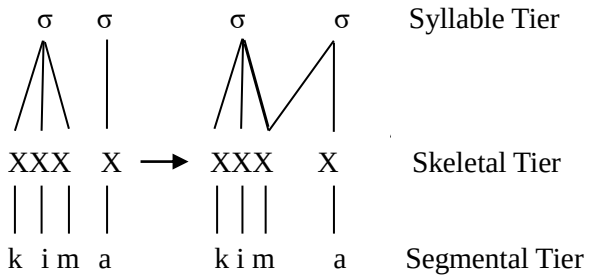



Figure 2. Ambisyllabicity rule

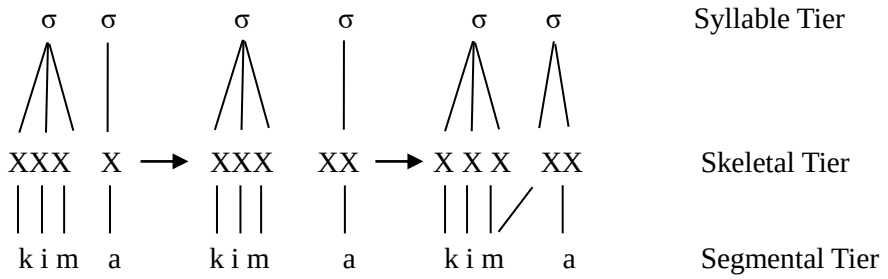



Figure 3. Gemination rule

As Figures 1, 2, and 3 indicate, the Skeletal Tier is the crucial level in the three-tier autosegmental representations and which represents the left-to-right order and duration of segments and is thus referred to as the

---

governed by a syllable can be treated as a single phonetic unit and linked to an X.

timing tier as well. Syllables are processed as individual units independent from segmental sequences at a separate tier, and link with the Skeletal Tier and the Segmental Tier. In other words, associations between every two levels can be different though the three levels are interconnected.

With the above discussion in mind, an intriguing question arises from the comparison between *kim a* 'gold' and *kim aʔ* 'golden duck': If the former is simply resyllabified or undergoes ambisyllabicity, the duration of its [m] on the surface is not predicted to be different from that of the latter, whose phonetic representation of [kim aʔ] 'golden duck' is realized autosegmentally in Figure 4. In other words, the duration of [m] in the phonetic representations of *kim a* 'gold' derived with resyllabification as [ki ma] (see Figure 1) and ambisyllabicity as [kima] (see Figure 2) is not predicted to be different from the same duration in [kim aʔ] 'golden duck'. The consonant [m] occupies one single X slot at the Skeleton Tier in all three representations.

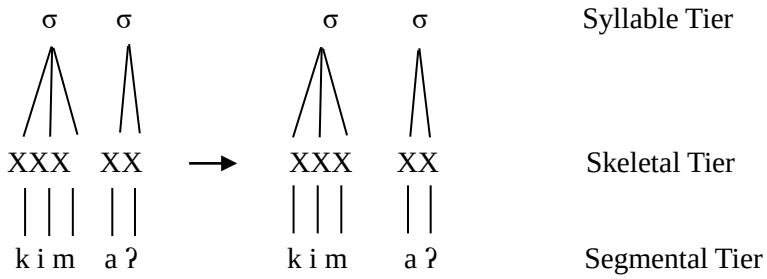



Figure 4. Derivational process of *kim aʔ* 'golden duck'

Likewise, we can compare the gemination of *kim a* 'gold' to that of *kim ma* 'place name'. The duration of [m] in the phonetic representation of [kimma] 'gold' derived via gemination is not predicted to be different from the same duration in the phonetic representation of [kimma] 'place name'. The derivational process of the latter is illustrated as follows:

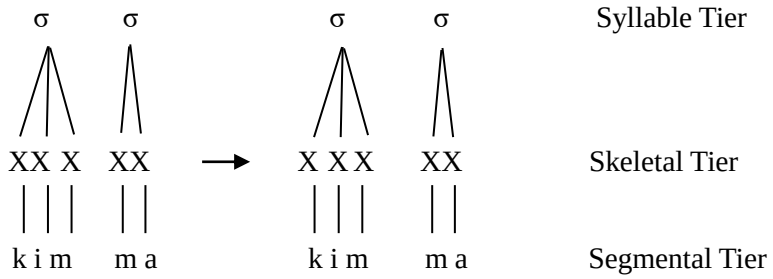



Figure 5. Derivational process of *kim ma* 'place name'

Comparing Figure 3 and Figure 5, we should notice that [m] occupies two X slots in the phonetic representation of [ki<u>mm</u>a] 'gold' derived with gemination, and the same duration of [m] occurs in the phonetic representation of the compound word [ki<u>mm</u>a] 'place name'; one cannot easily distinguish the two surface forms [kimma] 'gold' and [kimma] 'place name'.

To solve the problem caused by this mismatch among the duration of [m] in the three words, an account based on prosodic structure is proposed by Chiang (1992). In her study, Chiang investigates the duration of [m] in four Taiwanese words with different prosodic structures. The consonant [m], depending on its prosodic position, can be defined as a fake geminate in [kim][ma] 'place name', a true geminate in [kimma] 'gold', a single onset consonant in [k$^h$i][ma] 'at least', and a single coda consonant in [kim][aʔ] 'golden duck'. The results show that the durational difference is not significant between [k$^h$i][ma] 'at least' and [kim][aʔ] 'golden duck' but is significant for the other pairs: The fake geminate [m] in [kim][ma] 'place name' is the longest among the four, and the second longest is the true geminate [m] in [kimma] 'gold'. The single consonant [m] in [k$^h$i][ma] 'at least' and [kim][aʔ] 'golden duck' has the shortest duration (pp.69-83).

As the duration of [m] in phonetic forms varies, Chiang states that a desirable distinction in terms of the phonetic realization of [m] is needed, which lies behind some differences in prosodic structure. Words with a fake geminate and a single consonant should consist of two phonological words (i.e., [kim]PW [ma]PW 'place name'; [k$^h$i]PW [ma]PW 'at least';

[kim]PW [aʔ]PW 'golden duck'). By contrast, those with a true geminate are composed of a single phonological word.

Chiang's research is significant since she points out that the structural differences are the source of the durational differences. However, the duration of the [m] on the three surface forms is not predicted to differ for each of the forms. For example, both surface forms of [kimma] 'gold' and [kim][ma] 'place name' have two identical [m]s occupying the two timing slots (i.e., the equivalent numbers of [m]s); apparently "a desirable distinction" of durational differences of the two words would be hard to be justified. Furthermore, whether or not the transitional sound in [kimma] 'gold' is a geminate still requires more acoustic evidence. In regard to [kim][ma], it seems that this surface form is normally articulated in a continuous fashion; presumably there may be a link connecting the two phonological units. All in all, on top of Chiang's framework of structural differences, this present study attempts to propose phonetic realizations of the three lexical items, which are distinguishable in the duration of the three [m]s.

Along this line, this paper follows Chiang's (1992) experimental design to investigate durational variation of [m] across the three words *kim ma* 'place name', *kim a* 'gold', and *kim aʔ* 'golden duck'. However, in our study we increase the number of speakers in the recording session to 12 to rule out sociolinguistic factors that could influence phonetic outputs. In addition, we replace the phrase *k$^h$i ma* 'at least' with *kim aʔ* 'golden duck' for the following reasons: Firstly, it is to agree in terms of their lexical category (i.e., noun). Secondly, the underlying representations of the three target words share the identical segments /kima/, which also allow us to control the first syllable in our comparative analysis. Finally, since the duration of [m] in [k$^h$i ma] 'at least' and [kim aʔ] 'golden duck' does not vary notably, excluding the former one should not drastically affect the conclusion of our comparative analysis.

To recapitulate, our main research questions are as follows: First, is the transitional sound in *kim a* 'gold' a geminate? Second, is there any transitional sound in *kim ma* 'place name'? Third, is the durational difference in [m] among the three words perceivable for native speakers? Fourth, how do we transcribe the phonetic representations of the three

words? The above issues will further be discussed in section 3 of this paper.

## 2. METHODS

Our study defines the phonetic representations of *kim aʔ* 'golden duck', *kim a* 'gold', and *kim ma* 'place name' in accordance with the acoustic, articulatory, and perceptual data of native speakers. Six male and six female native speakers of Taiwanese were randomly recruited for the recording session.[6] The 12 participants spoke Taiwanese on a daily basis and did not report any reading or articulatory problems. After the recruitment, the research goal of the recording session was elaborated to the 12 participants, who then agreed that their production data would be used for research purposes only.

The materials in the recording session are composed of the three target words in the carrier sentence "I read ____ three times" written in Chinese characters. The order of the three sentences is illustrated in (1):

(1) Gua liam *kim aʔ* sãpien. 'I read golden duck three times.'
Gua liam *kim a* sãpien. 'I read gold three times.'
Gua liam *kim ma* sãpien. 'I read place name three times.'

The reason for embedding the target words in the carrier sentence is to control the duration of the target words; length variation emerging from an isolated context or different phrase positions can be avoided. The target words between /-m/ in /liam/ 'read' and /s-/ in /sã/ 'three' are easily segmented from the context. In our pilot study, we did not discover any mistakes of such design, although the target word *kim ma* was easily produced as *kim be* 'golden horse'.

The recording session proceeded in a professional recording studio, during which participants would practice reading the three sentences beforehand. If *kim ma* was produced as *kim be* 'golden horse',

[6] The participant numbers were balanced for each sex following Ladefoged's (2003) experience; he stated that six speakers for each sex would be enough for systematic differences between male and female speech.

participants were explicitly instructed to articulate *ma* as in the place name 'Matsoo', rather than *be* as in the pronunciation of 'golden horse'. When recording started, the participants had to repeat each Chinese written sentence three times at a speech rate as consistent as possible. After the recording session, each participant was asked to express an overall impression of the phonetic similarities and differences among the three target words in accordance with what he/she read in the three target stimuli. He/she could repeat the words as many times as they needed. Since the words were written in Chinese characters rather than the spelling or underlying representations, the participants' reaction to the phonemic forms would be reduced to a certain extent.

If the participants did not know how to respond, the interviewer would prompt them to make a simple phonetic comparison or contrast of each pair of words, i.e., *kim aʔ* vs. *kim a*, *kim a* vs. *kim ma*, or *kim ma* vs. *kim aʔ*. This self-expression process was recorded as well and the length for both the reading and self-expression sessions was approximately three minutes.

The analysis of the collected data was twofold: The production of target words in the carrier sentence was analyzed quantitatively, and the expression data were analyzed qualitatively. In the first analysis, the most stable sentence among the three attempts was selected, so the total number of tokens included in the analysis was 12 speakers × 3 target words = 36 tokens, which were extracted and processed using Praat 5.0.21 retrieved from http://www.praat.org. The scope of the phonetic analysis covers sound wave, intensity, formant, and frequency spectrum. The duration of [m] in the three target words was also measured and analyzed statistically; whether the segment length differs significantly in the three words was of primary analytical interest.

## 3. ANALYSIS AND DISCUSSION

Before moving to the analysis, our primary goal is repeated here: The transitional sounds of the three words are analyzed based on the acoustic properties of the words, and also based on native speakers' production

and perception. Following the analysis, this study compares representational differences at the phonetic level among the target words.

### 3.1 Acoustic Analysis

This section concentrates on an acoustic analysis that aims to examine whether the durational differences are statistically significant. We firstly extract and analyze the spectrogram of a sample token of 'golden duck' produced by one of our participants (see Figure 6):

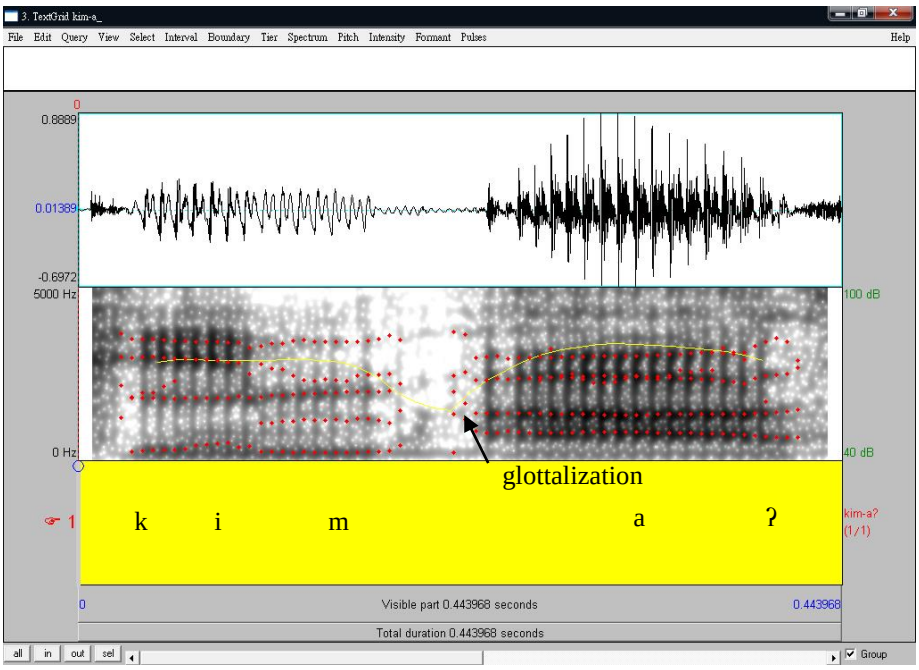



Figure 6. Spectrogram of [kim ˀaʔ]

The scope of the spectrogram covers the two syllables and the boundary in between in the word *kim aʔ*. The first syllable starts with the irregular noise [k], which is followed by the regular waves of the sonorants [i] and [m] toward the right edge of the syllable. The second

syllable carries a glottal stop that leads to a slightly pre-glottalized phenomenon. Due to its short duration and weak energy, glottalization is mostly visible in the section where intensity is low and the spectrogram becomes lighter as indicated by the arrow (intensity is represented with a curve in the figure). The regular waves of [a] follow the glottalization process and end with the presence of the glottal stop [ʔ]. On the right side of the wave form is the irregular noise [s]; namely, the onset consonant of the syllable [sã] 'three' in the carrier sentence 'I read golden duck three times'. Since no sound fills the boundary between the two syllables [kim] and [aʔ], the wave form is interrupted, the amplitude drops significantly, and the spectrum fades. That is, the sound vanishes at this point, and the gap is the boundary position in 'golden duck' (the silence gap in 'golden duck' is the boundary in this word).

After the above analysis of the spectrogram of 'golden duck', the spectrogram of 'gold' is represented in Figure 7. The spectrogram is also a sample token selected from the same speaker.

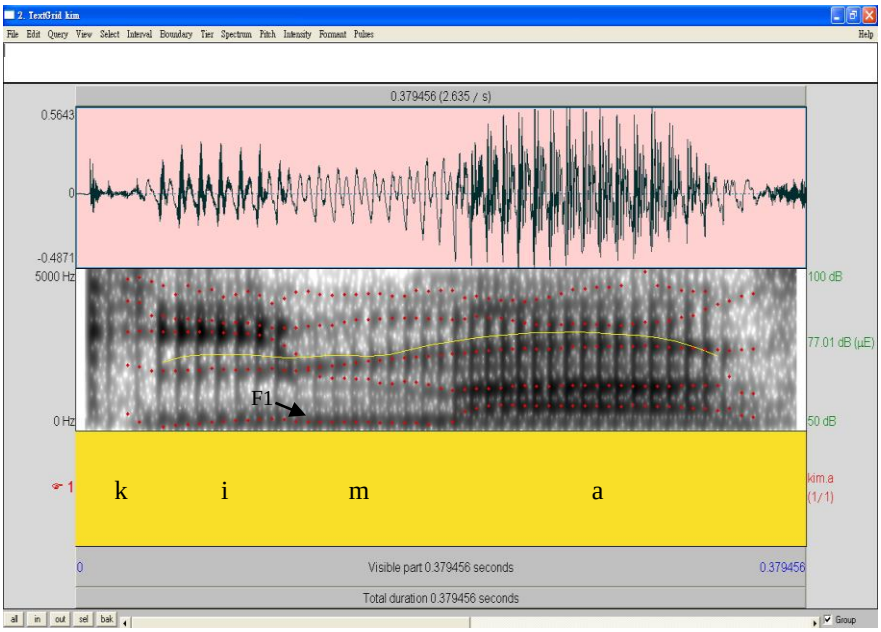



Figure 7. Spectrogram of [kim$^{ma}$a]

Similar to the spectrogram of *kim aʔ*, the spectrogram of *kim a* is also composed of two syllables with their boundary in between. In general, the sound waves of every segment in the two words are alike. One major mismatch, however, is that the boundary in *kim a* is occupied with a transitional sound, which is [ᵐᵃ] (the sound that fills in the boundary in *kim a* is the transitional sound [ᵐᵃ]). The sound wave of the transitional sound in the boundary does not vanish, the intensity does not drop, and the spectrum does not fade, either. The actual sound is thus continuous rather than intermittent (compare Figure 7 with Figure 6). When a sound is present at the boundary, we cannot define its range depending on the differences in sound wave, amplitude, or spectrum; the edges can only be discovered with recourse to F1 changes.

All sonorants have F1, which is represented with dots at the bottom of the figure. Although among sonorants, a nasal has an F1 band with a lighter color, it is still visually identifiable. F1 is inversely related to tongue height: The higher the tongue position is, the lower the F1 frequency will be, and vice versa. Since the [m] sound at the end of the first syllable has a higher tongue position, it has a lower F1 frequency (around 200 to 300 Hz). However, with a lower tongue height, the [a] sound at the beginning of the second syllable has a higher F1 frequency (around 700 to 800 Hz). We can thus predict a change from an F1 range 200-300 Hz to an F1 range 700-800 Hz within the transitional sound.

While changes in F2 are also a possible indicator, they are less observable. This is because F2 is related to tongue frontness, but tongue advancement or retraction is less obvious than tongue's lowering from [m] to [a]. Therefore, we solely rely on the more salient F1 changes in this paper. The spectrogram in Figure 7 can be illustrated in Figure 8 marked with syllable boundaries (acoustic information other than formants is excluded for easier visual inspection):

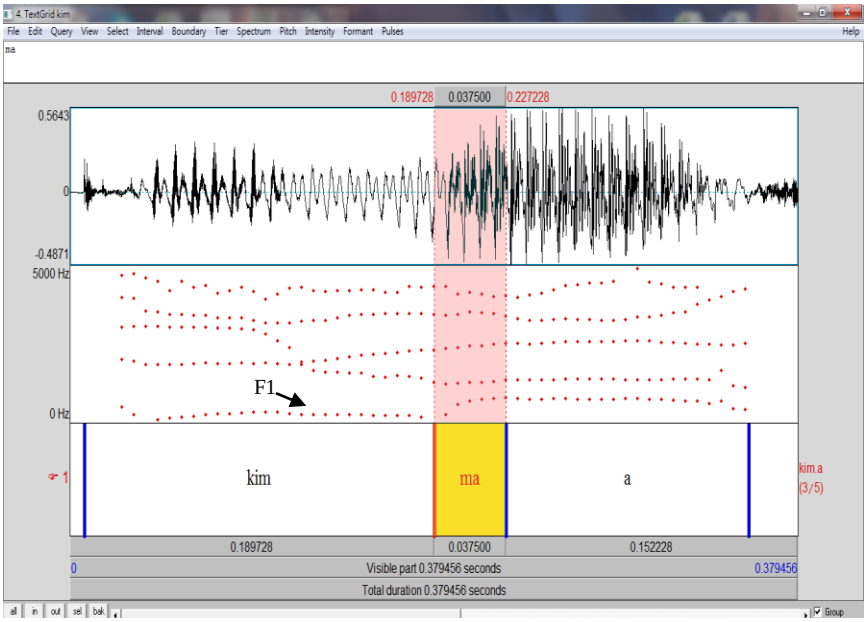



Figure 8. Syllable boundaries in [kim$^{ma}$a]

According to the fact that F1 is inversely related to tongue height, the range of the transitional sound in *kim a* can be defined as a section where the frequency of F1 drastically differs. For this speaker, the transitional sound corresponds to the section where the F1 range changes from 2-300 Hz to 6-700 Hz. The F1 frequencies at different time points are demonstrated in Appendix A, and are excerpted in Table 1:

Table 1. F1 frequencies at different time points in 'gold'[7]

| Time(s) | Frequency(Hz) | Frequency differences (Hz) |
|---|---|---|
| … | … | … |
| 0.1710 | 239.0636 | 1.2699 |
| 0.1772 | 234.2187 | -4.8449 |
| 0.1835 | 235.8215 | 1.6028 |
| 0.1897 | 246.5159 | 10.6943 |
| 0.1960 | 331.8794 | 85.3635 |
| 0.2022 | 493.2503 | 161.3709 |
| 0.2085 | 587.2845 | 94.0342 |
| 0.2147 | 634.8721 | 47.5876 |
| 0.2210 | 660.9060 | 26.0339 |
| 0.2272 | 687.6811 | 26.7751 |
| 0.2335 | 702.3803 | 14.6992 |
| 0.2397 | 705.6240 | 3.2437 |
| 0.2460 | 709.4479 | 3.8239 |
| 0.2522 | 704.2583 | -5.1897 |
| … | … | … |

Table 1 includes individual F1 frequencies at different time points between 0.1710s and 0.2522s and frequency differences between every two sequential time points. We can observe a trend of F1 changes in the column of frequency differences: Frequency differences are initially small and stable, which then increase remarkably before reducing to small values again. The six significant frequency differences are 85.3635 Hz, 161.3709 Hz, 94.0342 Hz, 47.5876 Hz, 26.0339 Hz, and 26.7751 Hz in the shaded cells.[8] The frequency difference in the first shaded cell (i.e., 85.3635 Hz) is the difference between the two F1 frequencies at

[7] All the values are rounded to the 4th decimal place.
[8] The significance of frequency differences is assumed as above 20 Hz in this paper.

0.1897s and 0.1960s, and the one in the sixth shaded cell (i.e., 26.7751 Hz) is the difference between two F1 frequencies at 0.2210s and 0.2272s. We can therefore conclude that the significant F1 changes occur between 0.1897s to 0.2272s; the area before 0.1897s is the first syllable [kim], and the one after 0.2272 is the second syllable [a]. Frequency differences in these two areas are relatively small and stable.

If we compare Figure 7 with Figure 8, we should recognize two segments involved in the transitional sound from 0.1897s to 0.2272s from the spectrogram of *kim a*. The spectrum in the first part of the transitional sound has a lighter color than in the second part of the sound. This is because the nasal sound [m] has weaker energy and formants in its spectrum than vowels. Furthermore, we may also see the trace of two segments from the waveform of the transitional sound: The waveform shape of the first half in the transitional area differs from the shape of the second half of the area. The first half has a waveform with larger intervals between cycles, which is similar to the waveform of [m], whereas the waveform in the second half has smaller intervals between cycles as in the waveform of [a]. The two above cues allow us to consider the transitional sound a sequence [$^{ma}$] composed of [m] and [a].

There is a notable difference between the transitional sound [$^{ma}$] and the syllable [ma] 'horse'. This transitional sound [$^{ma}$] (0.1897s~0.2272s, shorter duration of 0.0375s) is between the two syllables and has a large slope rate of F1. By contrast, the syllable [ma] (0.0071s~0.2246s, longer duration of 0.2175s, as shown in Figure 9) includes stable F1 frequencies of [m], a large slope rate of the tautosyllabic transitional sound [$^{ma}$], and is eventually followed by stable F1 frequencies of [a]. The spectrogram of the syllable [ma] 'horse' is demonstrated in Figure 9 and its F1 frequencies at different time points are quantified in Appendix A. As Figure 8, Figure 9, and Appendix A indicate, the heterosyllabic transitional sound [$^{ma}$] possesses a large slope rate, whose F1 is from 0.1897s to 0.2272s, whereas the syllable [ma] possesses three sections--a small slope rate of [m] (0.0071s~0.0394s), a large slope rate of the tautosyllabic transitional sound [$^{ma}$] (0.0394s~0.0644s), and a small slope rate of [a] (0.0644s~0.2246s). No doubt the heterosyllabic transitional sound [$^{ma}$] is much shorter than the syllable [ma], thus the former one is only considered a connection between the two syllables.

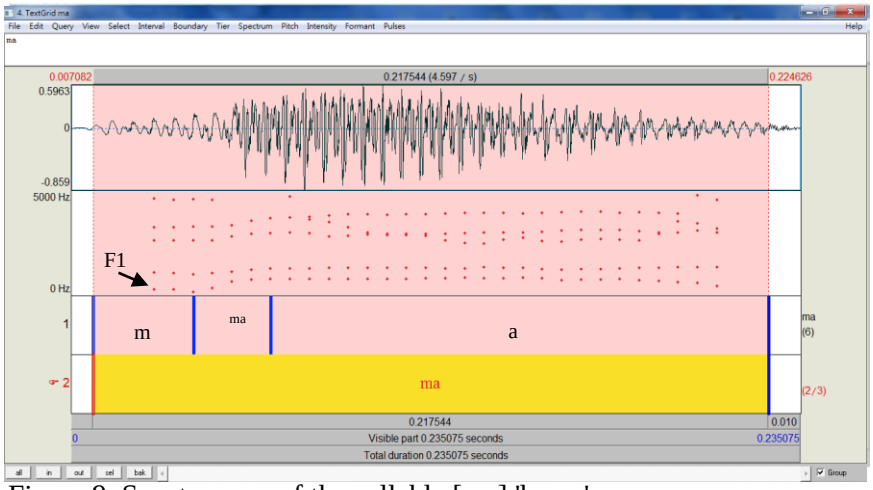


Figure 9. Spectrogram of the syllable [ma] 'horse'

Finally, we move to analyze the spectrogram of *kim ma* in Figure 10, which is likewise extracted from a sample token produced by the same speaker.

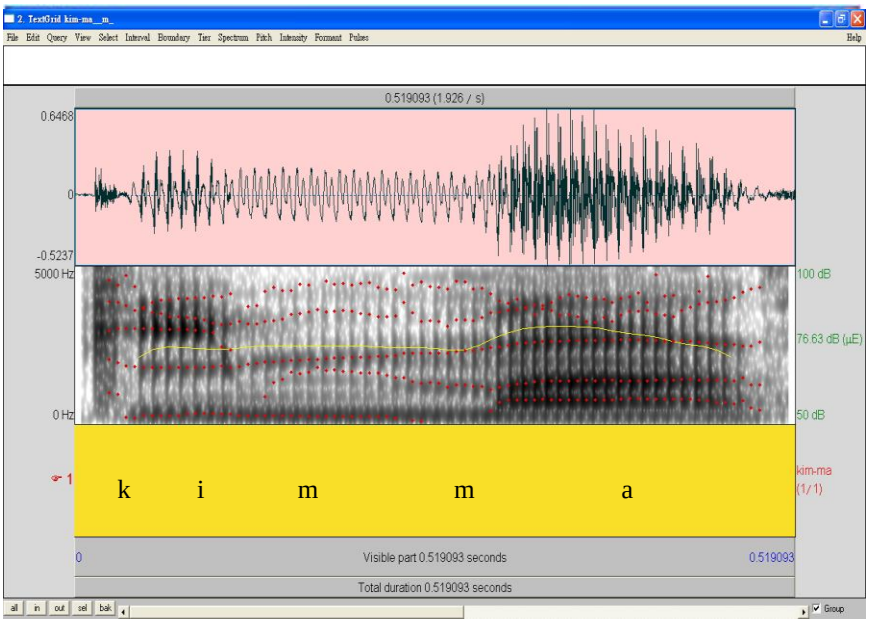


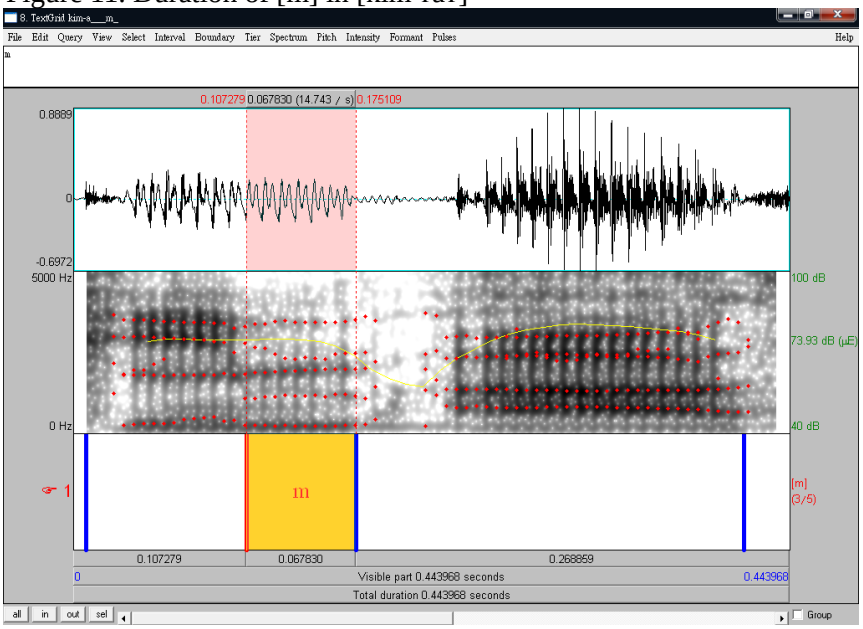

Figure 10. Spectrogram of [kim$^{mm}$ma]

Figure 10 is the spectrogram of *kim ma*, which is composed of two syllables with a boundary in between as in the spectrograms of *kim aʔ* and *kim a*. Generally speaking, the sound waves of every segment in the three words are approximately alike. However, the boundary of *kim ma* is occupied with a sound (the sound that fills the boundary in *kim ma* is the transitional sound [$^{mm}$], which will be described in 3.2 and 3.3), while the boundary of *kim aʔ* is empty. Although the boundaries in *kim ma* and *kim a* are filled with transitional sounds, the former has a longer duration of [m] than the latter.

With the spectrograms in Figures 6, 7, and 10, we can investigate whether the duration of [m] is significantly different across the three words with a statistical analysis, which can thus shed light on whether native speakers can solely rely on the duration of [m] to identify the three words. The duration of the nasal sound [m] can be determined acoustically with changes in its spectrum and waveform: In its spectrum, the color is relatively lighter for the nasal sound due to weak energy. The

waveform shape of the nasal sound also slightly differs from that of adjacent vowels, and the difference may help measure the duration of [m] as well. The duration of [m] in *kim aʔ* in Figure 6, *kim a* in Figure 7, and *kim ma* in Figure 10 can be located below, following the above guidelines:

Figure 11. Duration of [m] in [kim ʔaʔ]



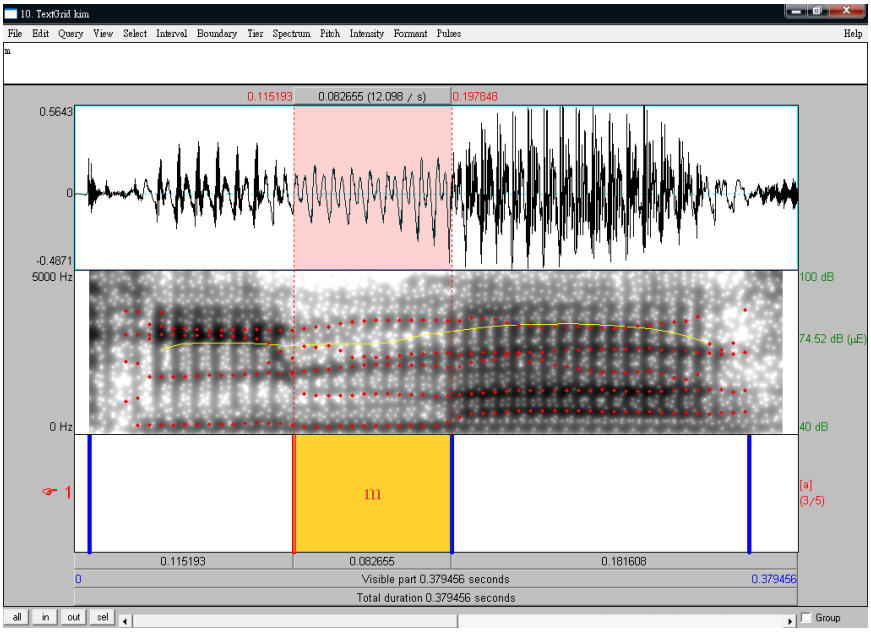



Figure 12. Duration of [m] in [kim$^{ma}$a]

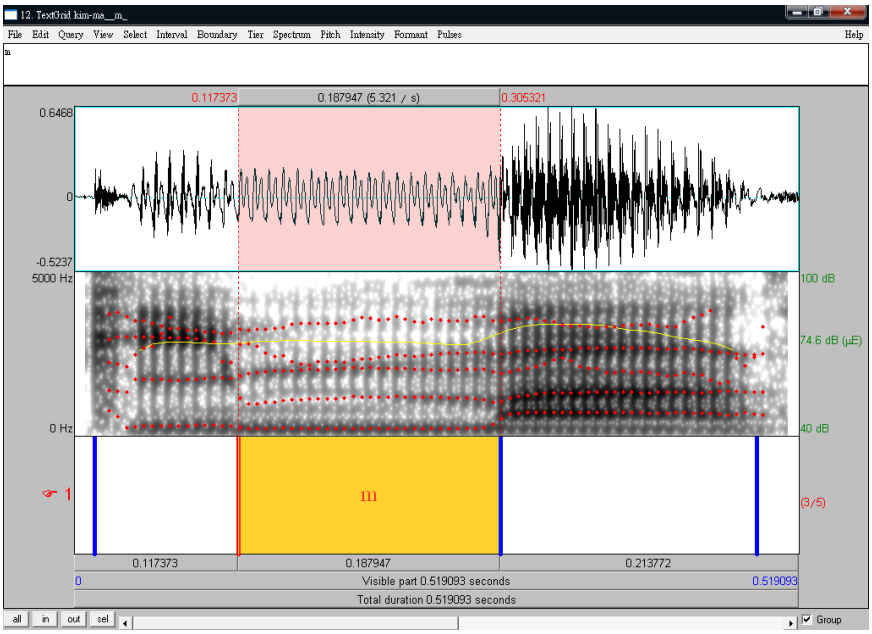



Figure 13. Duration of [m] in [kim$^{mm}$ma]

Figures 11, 12, and 13 show that the duration of [m] is 0.07s in *kim aʔ* (17% of the total duration 0.41s), 0.08s in *kim a* (23% of the total duration 0.35s), and 0.19s in *kim ma* (40% of the total duration 0.48s). If the proportion of [m] in *kim aʔ* (i.e., 17%) is treated as the reference level, the proportion ratio of [m] for the three words can be converted from 17:23:40 into 1.0:1.4:2.4. In this way, we can organize the quantitative data of our 12 speakers as in Table 2:

Table 2. Proportion ratios of [m] in the three surface forms

| | [kim ˀaʔ] | [kim$^{ma}$a] | [kim$^{mm}$ma] |
|---|---|---|---|
| M1 | 1 | 1.4 | 2.4 |
| M2 | 1 | 1.4 | 1.7 |
| M3 | 1 | 1.7 | 2.1 |
| M4 | 1 | 1.3 | 1.6 |
| M5 | 1 | 1.5 | 2.1 |
| M6 | 1 | 1.7 | 2.1 |
| AVG | 1 | 1.5 | 2 |
| F1 | 1 | 1.3 | 1.6 |
| F2 | 1 | 1.4 | 2.1 |
| F3 | 1 | 1.3 | 1.7 |
| F4 | 1 | 1.9 | 2.4 |
| F5 | 1 | 1.2 | 1.5 |
| F6 | 1 | 1.9 | 2.7 |
| AVG | 1 | 1.5 | 2 |
| Total AVG | 1 | 1.5 | 2 |

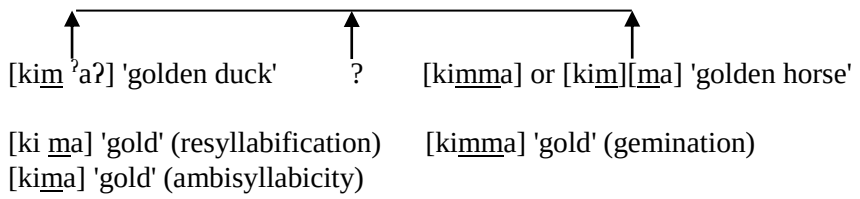
*Note*. M: Male; F: Female

Out of the 12 speakers in Table 2, six are male and six are female. The example tokens discussed in this section were selected from the data produced by M1. In Table 2, the numbers suggest that the average proportion ratio of [m] for the three words demonstrated by our male and female speakers is identical (i.e., 1:1.5:2). Therefore, there is no gender difference in terms of the production of [m] in the three Taiwanese words. The durational differences across the three words were analyzed using one-way ANOVA, which are significant ($F(2, 33) = 44.2$; $p < 0.001$). Furthermore, multiple comparisons also indicate a significant durational difference between every two words. Specifically, the difference in the duration of [m] is significant for the three pairs *kim aʔ* vs. *kim a*, *kim a* vs.

*kim ma*, and *kim ma* vs. *kim aʔ*. This finding is consistent with Chiang's (1992) results suggesting a significant relative difference in the duration of [m] among the three words *kim aʔ* (or *kʰi ma* 'at least'), *kim a*, and *kim ma*. We can therefore conclude that the durational differences, which native speakers use to distinguish the three words and are thus reflected in the words' phonetic representations, are statistically significant.

The significant variation in the duration of [m] across the three words can serve as a reference for us to examine the plausibility of various phonological rules in previous studies. We can compare the duration of [m] in *kim a* generated from resyllabification, ambisyllabicity, and gemination to that in *kim aʔ* and *kim ma* as in (2):

(2) Relative position of [m] based on durational differences

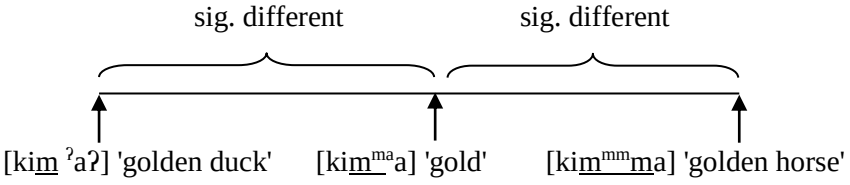



From illustration (2), we can see that since the two representations [ki ma] (via resyllabification) and [kima] (via ambisyllabicity) of 'gold' and the representation [kim ˀaʔ] of 'golden duck' have only one [m], they occur in the same position and do not have any obvious difference. Similarly, the representation [kimma] of 'gold' derived via gemination has two [m]s as in the representation [kimma] or [kim][ma] of 'place name'; there is no notable representational differences since these representations occur in the same position as well.

Based on the proportion ratio 1:1.5:2 calculated above, the [m] sound derived via some phonological processes in the phonetic representation 'gold' should occur between the duration of [m] in 'golden duck' and that in 'place name' to specify the representational differences among the three words. As the above three phonological rules fail to address the significant durational differences, we should seek an alternative

explanation with recourse to the concept of transitional sound proposed in this paper as demonstrated in (3):

(3) Relative position of [m] based on durational differences




In (3), the three phonetic representations corresponding to the three words are [kim ʔaʔ], [kim$^{ma}$a], and [kim$^{mm}$ma], respectively, and we should be able to notice the representational differences of [m] among the three words. That is, the form [kim ʔaʔ] has one single [m], the form [kim$^{ma}$a] has one [m] and one [$^{m}$], and the form [kim$^{mm}$ma] has two [m] sounds and two [$^{m}$] sounds. The phonetic form [kim$^{ma}$a] rather than [ki ma], [kima], or [kimma], is more likely to be located in the position between [kim ʔaʔ] and [kim$^{mm}$ma], so the significant difference can be specified. Perhaps the duration of more than 1 [m] and less than 2 [m]s of [kim$^{ma}$a] may provide one possible explanation for why native speakers always hesitate in deciding whether it is purely one [m] or two [m]s for the realization of this word. This speculation may rely on the experimental discussions of Kao (2003), Wang and Kao (2004), and Wang and Liu (2010) as well.

The researchers (Kao 2003; Wang and Kao 2004) conducted two experiments to examine the phonological status of the onset consonant in the vowel-initial suffix. In the first experiment (Syllable Inversion Experiment), the subjects were asked to reverse the order of bisyllabic *-a* attached words (e.g., *kim-a* is treated as *a-kim* or *ma-kim*), and the result significantly indicated that the onsetless *-a* was the preference. In the second experiment (Concept Formation Experiment), the subjects were trained to make a positive response to some target sounds, and then they were asked to tell whether they heard a consonant appear before the suffix *-a*. The results of the second experiment revealed that the positive and negative responses to the onset position of the suffix *-a* were

inconsistent -- some final consonants induced more onsetless responses to the suffix -*a*, while others did not.

Based on the results of the two experiments, the researchers concluded that the onset was recognized by the native speakers, yet had not been phonologized at the underlying level. Accordingly, they predicted that these liaison consonants were on the road to be fossilized into part of the suffix since the awareness of the onset was the precursor to language change. Aside from this conclusion, the phonetic duration of a final coda may be regarded as wavering between 1 unit and 2 units, i.e., 1[m] < [ki$\underline{\text{m}}^{\text{ma}}$a] < 2[m]s. Perhaps the findings of the two experiments to some extent may coincide with what I propose in this present paper.

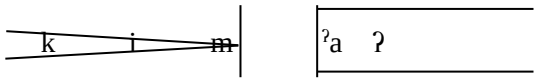
Except for the discrepancies in the surface forms among the three words, the structural differences among the three forms are noticeable and are considered as the source of the surface disparity. Since the four syllables of the two words *kim aʔ* and *kim ma* are free morphemes, they are emphasized altogether. By contrast, the word *kim a* is a suffixed word and its suffix is short and indistinct: This is perhaps because the suffix [a] is a bound morpheme and must be attached to a preceding stem. In other words, the two morphemes are more inclined to coalesce into one unit in ‘gold’ than in compound words. The phonetic realizations of the three words therefore can further be transcribed as [kim][ˀaʔ] 'golden duck', [kim$^{\text{ma}}$a] 'gold', and [kim]$^{\text{mm}}$[ma] 'place name'. As a result, the formal mechanism in deciding the existence/absence of the transitional slots in phonetic form would be: the transitional sound occurs in one prosodic word (e.g., [kim$^{\text{ma}}$a] 'gold') rather than two prosodic words (e.g., [kim][ˀaʔ] 'golden duck', [san][baʔ] 'lean meat', [k$^{\text{h}}$i][ma] 'at least') unless the coda and the onset are the same nasals, which are apt to induce connections (e.g., [kim]$^{\text{mm}}$[ma] 'place name', [sin]$^{\text{nn}}$[niu] 'bride', [tiɔŋ]$^{\text{ŋŋ}}$[n̩] 'middle').

### 3.2 Articulation

The foregoing discussion explained the similarities and differences in the phonetic representations of the three target words based on acoustic analyses. This section continues to provide a comparative analysis but seeks to demonstrate representational similarities and differences from

an articulatory perspective. First, the focus is on the articulation of *kim aʔ* 'golden duck', which is a compound word composed of two syllables (i.e., first syllable = [kim], second syllable = [ˀaʔ] or morphemes). During the production of [kim], speakers begin with a dorsal consonant [k], which is realized with a slightly open mouth. A vowel [i] immediately follows the consonant with a reduced oral cavity space and precedes a bilabial consonant [m] with labial closure at the end of the first syllable. The second syllable starts with a vowel [a] produced with a maximized oral cavity space and vocal cord vibrations. The chronological articulatory changes in the production of the two syllables are illustrated in (4):

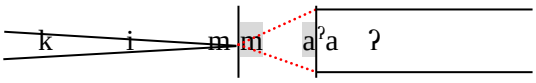
(4) Chronological articulatory changes in [kim] and [ˀaʔ]




From (4), it should be obvious that the space of the oral cavity gradually reduces to a complete closure in the first syllable, while the open space remains constant throughout the entire second syllable. The two vertical lines stand for the syllable boundaries between the two syllables. When the two syllables are combined as one single word, there is a continuous articulatory process from [k] in the first syllable to the glottal stop [ʔ] in the second syllable. Nevertheless, an articulatory gap appears at the boundary between the two syllables [kim] and [ˀaʔ], namely a silence area between the two morphemes. It is possible to measure the duration of the silence area, which, albeit inaudible, covers articulatory movements. The primary movement in this area is a gradual lip opening movement from [m] in the first syllable, which proceeds until the point before reaching the maximal oral cavity space required by the vowel [a] in the second syllable (glottalization occurs shortly before the beginning of [ˀaʔ]). This articulatory movement is represented with subscribed [$_{ma}$] to specify its difference relative to [m] with a complete lip closure and [a] with a maximal oral cavity space. The duration of the silence area is represented with dotted lines in (5). As the inaudible lip

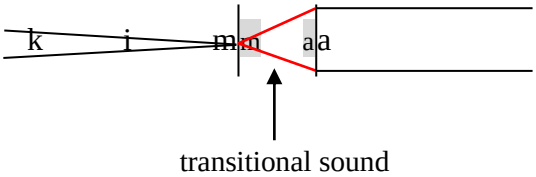
movement occurs at the articulatory gap in the middle of 'golden duck', we transcribe this area with a space, which turns the word's phonetic representation into [kim][ˀaʔ].

(5) Chronological articulatory changes in 'duck'



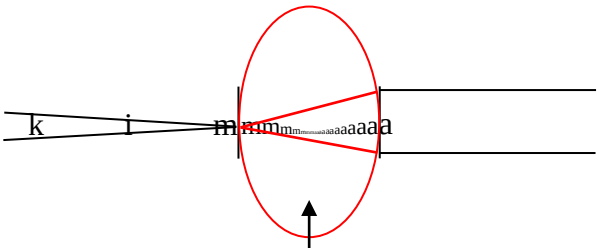

Now we shift our focus to the production of the suffixed form *kim a* 'gold', which is a combination of a free morpheme /kim/ and a bound morpheme /a/. The adjustment of the oral cavity space in *kim a* is similar to that in *kim aʔ*, except that there is no articulatory gap between the two syllables [kim] and [a]. On the contrary, a transitional sound occupies the boundary position to connect the syllables [kim] and [a]. The duration of this audible [ma] is represented with solid lines in (6), and the phonetic representation of 'gold' can be transcribed as [kimmaa]. The transitional sound is superscripted as in [kim$^{ma}$a] to (i) avoid the confusion that the transitional sound is visually akin to part of the preceding or following syllable, and (ii) underscore the fact that the transitional sound, unlike the surrounding syllables, has much shorter duration.

(6) Chronological articulatory changes in 'gold'




According to the adjustment made to the oral cavity, it is plausible to define the phonetic representation of *kim a* as [kim$^{ma}$a]. Illustration (7) shows the detailed transitional changes made by lengthening the transitional sound section in (6).

(7) Detailed changes of the transitional sound [$^{ma}$] in 'gold'



The gradual changes in the transitional sound [$^{ma}$] can be directly observed in illustration (7). The variation in the oral cavity space during the production of the transitional sound is represented with two solid lines 「<」, and the transition between the two segments during the production is represented with 「mmmmmaaaaa」. The onset of the transitional sound [$^{ma}$] is the release of the lip closure, and the mouth opening gradually increases toward the production of [a]. As the sound of [m] fades out, the vague sound of [a] is introduced and gradually becomes more clear; the mouth opening is enlarged until the movement reaches the articulatory target of [a]. In other words, the position of lips and tongue changes within the transitional sound, which involves the two segments [m] and [a]. Without the above transition, it would be hardly understandable why the [m] sound at the end of the first syllable is connected with the [a] sound at the beginning of the second syllable with no intervening interruption. Therefore, the so-called 'liaison' or 'linking' in previous literature is dubbed 'transitional sound' to manifest its transitional quality.

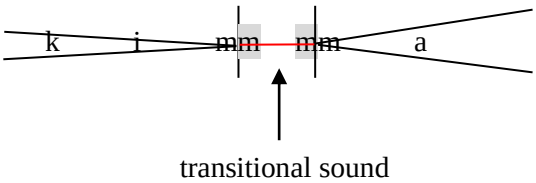
The compound word *kim ma* comprises two free morphemes /kim/ and /ma/, and the change in the oral cavity space during its production is similar to that in 'golden duck' and 'gold'. Producing the first syllable /kim/ in isolation results in a gradual oral closure as the articulatory movements travel through the sounds [k], [i], and finally [m]. The second syllable /ma/ starts with a lip closure [m], which is followed by a vowel [a] requiring a maximal oral cavity space; the oral cavity space

gradually increases to complete the [ma] sequence. The oral cavity changes of the two independent syllables [kim] and [ma] are illustrated in (8):

(8) Oral cavity changes during the production of [kim] and [ma]


If the two words 'kim' and 'ma' in the compound word *kim ma* are read separately as in (8), a gap occurs at the boundary between 'kim' and 'ma'. On the other hand, the production of the two syllables in connected speech is usually uninterrupted. The reason for the continuous production is the presence of a sound that fills the boundary between the two morphemes as a transition between the two syllables, which can be found in 'gold'; as a result, the transitional sound will occur between the coda consonant [-m] and the onset consonant [m-]. We thus treat the transitional sound in 'place name' as [$^{mm}$] and the phonetic representation of the word as [kim]$^{mm}$[ma]. The oral cavity changes during the production of 'place name' are represented in (9):

(9) Oral cavity changes during the production of 'place name'




From (9), it is notable that the transitional sound (solid lines) persists for a while and its length connects the two morphemes [kim] and [ma]. If the transitional sound is silent, the compound word 'place name' might be perceived as two independently produced words. The transitional sound [$^{mm}$] in 'place name', similar to the transitional sound [$^{ma}$] in 'gold', is

derived from the preceding and following [m] via bidirectional assimilation.

The proposal of the transitional sound [$^{mm}$] can account for the difference in phonetic representations between the words 'place name' and 'gold'. That is, the duration of the [m] sound is longer in [kim]$^{mm}$[ma] 'place name' than in [kim$^{ma}$a] 'gold'; [kim]$^{mm}$[ma] has two [m] sounds from the free morphemes and two [$^{m}$] sounds from the transitional sounds, whereas [kim$^{ma}$a] only has one [m] sound from the free morpheme and another from its transitional sound. It should thus be straightforward that the [m] sound is longer in 'place name' than in 'gold', and native speakers may simply rely on the durational difference to efficiently distinguish the two words from each other.

### 3.3 Perception

#### 3.3.1 Self-expression

Based on the self-expressions made by the 12 speakers, the similarities and differences among the three target words are summarized as follows:

(10) a. In *kim aʔ* and *kim ma*, both syllables are emphasized, whereas only the first syllable of *kim a* is emphasized and the suffix is not salient.

b. The duration of the consonant [m] is the longest in *kim ma* and the shortest in *kim aʔ*.

c. The word *kim a* has more [a] sound, whereas the word *kim ma* has more [m] sound.

d. During the production of *kim aʔ* and *kim a*, there seems to be a single lip closure gesture, whereas two lip closure gestures are seemingly realized when producing *kim ma*.

Some conclusions can be generalized from the participants' inter-word comparisons from (10a)-(10d). With (10a), we can conclude that *kim aʔ* and *kim ma* have the same compound word structure, which are perceived as [kim]+[$^{ʔ}$aʔ] and [kim]+[ma]. In particular, since the four

syllables are free morphemes, they are all emphasized. By contrast, the descriptions claim that only [kim] in *kim a* is emphasized, and the second syllable [a] is less prominent: It is because the suffix [a] is a bound morpheme and must be attached to a preceding stem. Again, the two morphemes seem to coalesce into "one unit" in *kim a* than in compound words.[9]

The descriptions in (10b) and (10c) suggest that our participants distinguished the three target words with the duration of [m], which is the longest in *kim ma* and the shortest in *kim aʔ*. According to (10c), they might have relied on the duration of [a] in identifying *kim a* and *kim ma* as well since [a] is longer in the former and [m] is longer in the latter. The above differences should therefore be expressed in the phonetic representations of the three target forms, which crucially lie in the variation of [m] and [a].

Finally, the self-expression in (10d) suggests that the production of *kim aʔ* and *kim a* only involves a single lip closure gesture, whereas there are two lip closure gestures in *kim ma*; apparently, the closure occurs in the production of the actual [m] sound. We can further assume that although only one lip closure gesture occurs in *kim aʔ* and *kim a*, there would be some difference between the two words, which has been revealed by the native speakers' self-expressions. In *kim aʔ*, the offset of [m] aligns with the right boundary of the first syllable, whereas in *kim a* the production of [m] continues after the end of the first syllable. The extra portion of [m] in *kim a* nevertheless is not sufficiently lengthened to be realized as one single [m] segment, or the result would be a form that cannot be distinguished from *kim ma*. Besides, during the production of *kim ma*, the two lip closure gestures are not interrupted with a silence gap. Instead, a slightly brief sound exists between the two gestures, which is not part of the vowel [i] or [a]. The sound should be extended from [m], whose duration would be observable, and this may also be one of the reasons why [m] in *kim ma* seems the longest.

---

[9] The word *kim a* 'gold' has a diminutive semantic structure, whereas both *kim aʔ* 'golden duck' and *kim ma* 'place name' possess lexical meanings.

### 3.3.2 Perceptual judgment

In this section we design a follow-up perceptual judgment experiment to investigate whether native listeners can differentiate the two words *kim a* and *kim ma* by means of ABX paradigm (McGuire 2010).[10] The voice samples consisted of each speaker's two target words edited from the two embedded sentences. The overall voice samples for the 12 speakers were 48 (12 speakers × 4 orders, ABA, ABB, BAB, BAA), for example, /kim a/-/kim ma/-/kim a/ or /kim ma/-/kim a/-/kim ma/. The balancing of AB ordering is to avoid recency effects. As to the intensity of the three stimuli, it was scaled to 70 dB. The interstimulus interval (ISI) was 500 ms. The order of the 48 ABX voice samples was randomized using Excel 2003.

59 qualified native Taiwanese listeners (14 male, 45 female) participated in the study. None of the participants reported any hearing difficulties. The questionnaire of ABX discrimination (see Appendix B) was distributed to every listener and the written instructions in the questionnaire directed the participants on how to respond to the stimuli. They were informed that the three "sounds" rather than the three "words" from each speaker were to be heard, so the awareness of underlying representations or orthography would be reduced to a certain degree. The three stimuli were presented in a series and the listener had to judge whether the third stimulus X was identical with A or B by marking their answers on the questionnaire. After a 3-trial practice, participants heard one speaker's ABX voice sample twice and they had 3000 ms to respond before the next voice sample began. It took approximately 15 minutes to judge the 48 samples. Results of the discrimination accuracy are shown in Figure 14 and Appendix C.

---

[10] As Footnote 1 indicates, the tone of *kim aʔ*$^{\underline{31}}$ is quite different from those of the other two words. In addition, the glottalization of the word *kim aʔ* can easily be perceived by native speakers as well. For these reasons, the word *kim aʔ* is removed from the experiment.

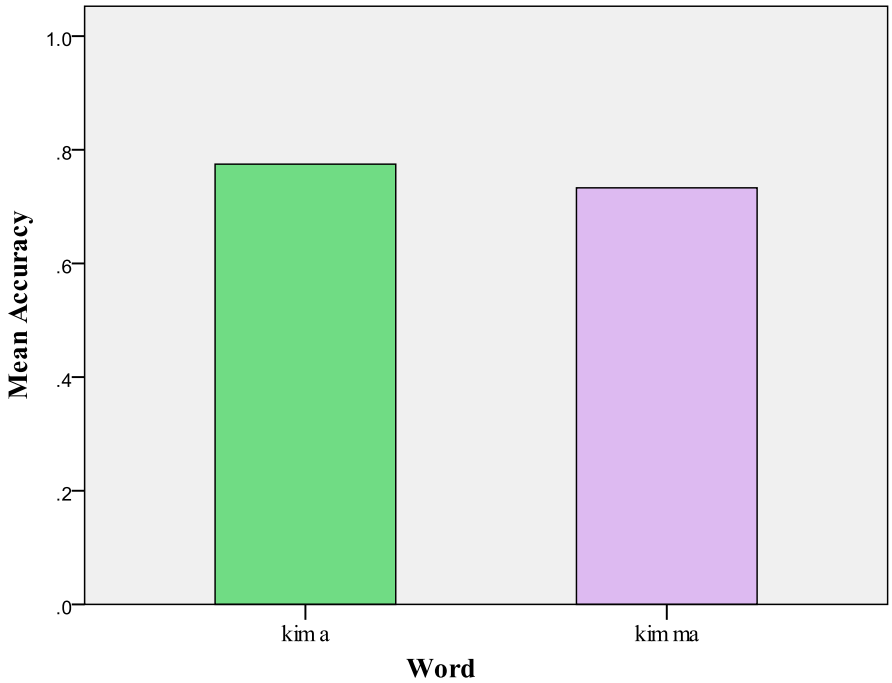



Figure 14. Discrimination accuracy of *kim a* and *kim ma*

As Figure 14 indicates, the native Taiwanese listeners' discrimination is equally good for *kim a* (mean accuracy = 0.77) and *kim ma* (mean accuracy = 0.73), which both surpass 0.50. Specifically, their discrimination of the contrast between the two words is notable, inferring that this contrast has become established in their perceptual system. The evidence of this perceptual discrimination task, therefore, supports our acoustic findings.

### 3.4 Summary

Taking evidence from acoustic analysis (3.1), articulation (3.2), and perception (3.3) into account, the autosegmental representations of the three target words can be proposed as follows:

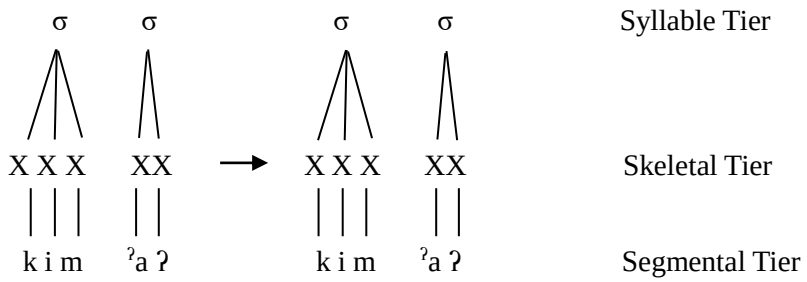



Figure 15. Derivational process of *kim aʔ* 'golden duck'

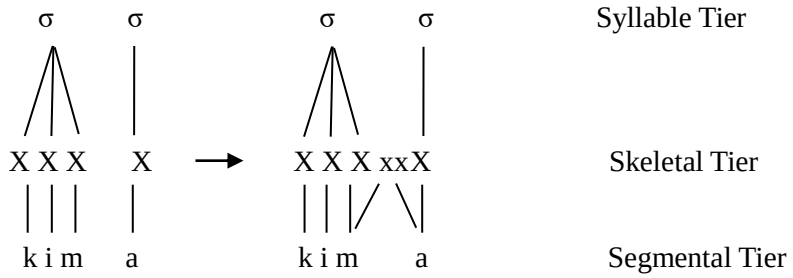



Figure 16. Derivational process of *kim a* 'gold'

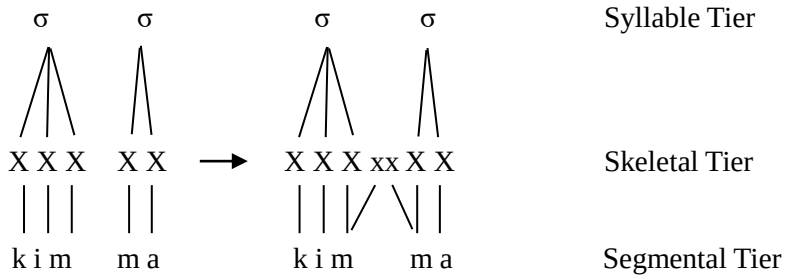



Figure 17. Derivational process of *kim ma* 'place name'

In Figure 15, there is a non-syllabic boundary in the middle of the Syllable Tier in [kim ˀaʔ], which is not filled with any sound; the sound sequence is interrupted at its midpoint. Therefore, only one single [m] sound is linked to an empty X slot. Since ‘golden duck’ is a compound word, its ultimate phonetic representation is [kim][ˀaʔ].

In Figure 16, the derivational process of [kim$^{ma}$a] indicates a non-syllabic transitional sound [$^{ma}$] that occupies the boundary, connects the preceding and following syllables, and therefore makes the sound sequence unbroken. This transitional sound [$^{ma}$] is derived via bidirectional spreading from [m] in the stem and [a] in the suffix, which fills the two previously preserved empty X slots (i.e., xx-slots) at the boundary between the two syllables, and this phonological process is dubbed bidirectional assimilation in this paper. According to this rule, the stem-final [m] in *kim a* is linked to an empty X slot, and the derived non-syllabic transitional sound [$^{ma}$] is associated with the empty xx slots. Since ‘gold’ is a suffixed form, its ultimate phonetic representation is [kim$^{ma}$a].

The reasons for the above transitional sound to be non-syllabic are twofold: First, the transitional sound occurs between two syllables at the surface level and thus never exists in underlying forms. The primary function of the transitional sound is to fill the cross-syllable boundary, rather than create an independent syllable. Second, it would be difficult to describe the transitional sound directly since it is realized with an extremely short length which cannot form a single syllable per se. For the above reasons, we transcribe the non-syllabic transitional sound as a superscripted sequence [$^{ma}$]. The X slots corresponding to the non-syllabic sequence on the Skeleton Tier are marked with lower-case symbols [xx], and those corresponding to the syllable [ma] are marked with capitalized symbols [XX].

Finally, the derivational process of *kim ma* in Figure 17 includes the non-syllabic transitional sound [$^{mm}$], which is also derived via bidirectional assimilation. Therefore, in *kim ma*, two [m] sounds are associated with the corresponding X slots, and the two derived [$^{m}$] sounds are linked to the xx slots. Note that the ‘place name’ is a compound word, and the morphological structure leads to its ultimate phonetic representation [kim]$^{mm}$[ma]. We adapt the X-Tier model rather

than the CV model in this paper due to different types of transitional sounds: The transitional sound [$^{ma}$] in [kim$^{ma}$a] is a CV sequence and the transitional sound [$^{mm}$] in [kim]$^{mm}$[ma] is a CC sequence. Due to the inconsistency, these transitional sounds are marked with X slots for simplification; x-slots do not intrinsically represent C or V, and the segmental identity of an X slot may change depending on the context.

Last but not least, our analysis leads to some intriguing questions: If [kim][ˀaʔ] and [kim]$^{mm}$[ma] are both compound words, why does bidirectional spreading not apply to the former to derive a transitional sound? Furthermore, why is a transitional sound absent in [kim][ˀaʔ] but present in [kim$^{ma}$a] when the stem [kim] immediately precedes a vowel in both words? One of the possible explanations might be an underlying prevocalic glottal stop, which is surfaced to block bidirectional spreading. In [kim]$^{mm}$[ma], there is no prevocalic glottalization with the [m-m] sequence to prevent the emergence of the transitional sound [$^{mm}$]. As to the form [kim$^{ma}$a], it can be treated as one single morphological unit as its components are more likely to be coalesced. Therefore, a transitional sound emerges as the merging process cannot complete without deleting prevocalic glottalization (i.e., glottal stop deletion). For example, to turn the word 'duckling' *aʔ a* into "one single unit", the stem-final glottal stop and the underlying initial glottal stop must be deleted first to derive a transitional sound. The primary difference in the case of [kim][ˀaʔ] is that on the one hand the morpheme *aʔ* in the word is underlyingly represented as /ˀaʔ/; on the other hand, this word consists of two free morphemes (i.e., a weaker merger pressure with 2 PrWd) as its structure indicates. In the end, the glottal stop deletion does not apply, and the prevocalic glottal stop intervenes between [m] and [a] and finally derives the form of [kim][ˀaʔ].

## 4. CONCLUSION

This current research explores the phonetic differences among the three Taiwanese words *kim aʔ*, *kim a*, and *kim ma* step-by-step in terms of their qualitative and quantitative analyses. Our conclusions are based on the acoustic analysis and native speakers' production & perception of the duration of [m] in these three words. As a result, the research questions are summarized as follows: (1) Our research suggests the phonetic representation of the suffixed form *kim a* to be [kim$^{ma}$a] since after the application of the bidirectional assimilation rule, it can be effectively distinguished from that of [kim][$^{ʔ}$aʔ] and [kim]$^{mm}$[ma]. (2) The durational difference of [m] across the three words is statistically significant, which is consistent with the production and perceptual judgment made by the native speakers. Therefore, the differences need to be expressed in the phonetic forms of the three words. In summary, the phonological structures corresponding to *kim aʔ*, *kim a*, and *kim ma* are /kim//$^{ʔ}$aʔ/, /kim a/, and /kim//ma/ respectively, and their corresponding derived phonetic structures are [kim][$^{ʔ}$aʔ], [kim$^{ma}$a], and [kim]$^{mm}$[ma].

The theoretical implication of this study perhaps can be scrutinized from the contrast of consonant length. In the literature, the consonant and vowel length contrasts have been attributed to the moraic structure, in which geminate consonants are distinguished from single consonants by assigning them a single underlying mora (e.g., Broselow et al. 1997; Hayes 1989). In this case, the word *kim a* would obtain three moras on the surface form [kim(μμ)ma(μ)], which has the same weight as *kim ma* [kim(μμ)][ma(μ)]. Such analysis violates our acoustic findings, in which the weight of *kim a* is "lighter" than the weight of *kim ma*. The moraic theory may be applicable to only the contrast of [CV(μ)][$\underline{C}$V(μ)] (Japanese *kita* 'arrived') vs. [CV$\underline{C}$(μ)(μ)][$\underline{C}$V(μ)] (*kitta* 'sliced'), but limited to the contrast of [CV(μ)][$\underline{C}$V(μ)] vs.[CV$\underline{C}$(μ)(μ)][V(μ)] (weight by position) vs. [CV$\underline{C}$(μ)(μ)][$\underline{C}$V(μ)] or [CV(μ)][$\underline{C}$V(μ)] vs. [CV$\underline{C}$(μ)(μ)$\underline{C}$V(μ)] (suffixed words) vs. [CV$\underline{C}$(μ)(μ)][$\underline{C}$V(μ)]. In my proposal, I state that the length differences are attributed to phonetic transition across different prosodic boundaries. The durational contrast of the consonant [m] therefore can explicitly be observed from the surface forms of [kim][$^{ʔ}$aʔ] vs. [kim$^{ma}$a] vs. [kim]$^{mm}$[ma]. Indeed, the Taiwanese

case is distinguished from the previous cases and should be captured differently.

In our future study, we can extend the concept of two-slot transitional sounds to explain similar structures in Taiwanese. The application and extension of our two-slot framework will allow a re-examination of the transitional sound in Chinese, English or any other language. The two-slot transitional sound still requires cross-linguistic explorations to eventually elucidate the features shared by human languages.

## REFERENCES

Ang, Ui-Jin. 1985. *Studies on the Tones of Taiwan Holo*. Taipei:The Zhili Press.

Broselow, Ellen, Su-I Chen, and Marie Huffman. 1997. Syllable weight:convergence of phonology and phonetics. *Phonology* 14.1:47-82.

Cheng, Liang-Wei Robert. 1997. *Taiwanese and Mandarin Structures and Their Developmental Trends in Taiwan I: Taiwanese Phonology and Morphology*. Taipei:The Yuanliou Publishing Co., Ltd.

Cheng, Liang-Wei Robert, and Shu-Juan Cheng Xie. 1977. *Phonological Structure and Romanization of Taiwanese Hokkien*. Taipei:The Student Book Company.

Chiang, Wen-Yu. 1992. *The Prosodic Morphology and Phonology of Affixation in Taiwanese and Other Chinese Languages*. University of Delaware. Doctoral dissertation.

Chung, Raung-Fu. 1996. *The Segmental Phonology of Southern Min in Taiwan*. Taipei: The Crane Publishing.

Clements, George, and Samuel Keyser. 1983. *CV Phonology: A Generative Theory of the Syllable*. Cambridge, MA: MIT Press.

Goldsmith, John. 1976. *Autosegmental Phonology*. MIT.Doctoral dissertation.

Hayes, Bruce. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20:253-306.

Kao, Shih-Ting. 2003. *The Reality of Geminated Consonants between Stem and Suffix in Taiwan Southern Min*. MA thesis, National Tsing Hua University.

Ladefoged, Peter. 2003. *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*. MA:Blackwell.

Levin, Juliette. 1985. *A Metrical Theory of Syllabicity*. MIT. Doctoral dissertation.

Lin, Ping-You. 2006. *A Study of the Taiwanese Suffix -a$^{53}$: An Integrated Autosegmental Analysis*. De La Salle University. Doctoral dissertation.

Lin, Ping-You. 2012a. A study of Taiwanese suffixed words with syllable-final stops. *NTU Studies in Taiwan Literature* 12:123-146.

Lin, Ping-You. 2012b. The transition sound of the Taiwanese suffixation word *kim a* 'gold'. *Journal of Taiwanese Languages and Literature* 7.2:55-86.

Lin, Yen-Hwei. 1989. *An Autosegmental Treatment of Chinese Segments*. University of Texas, Austin. Doctoral dissertation.

Lin, Yen-Hwei. 2004. Chinese affixal phonology: Some analytical and theoretical issues. *Language and Linguistics* 5.4:1019-1046.

McGuire, Grant. 2010. A brief primer on experimental designs for speech perception research. Retrieved from (http://people.ucsc.edu/~gmcguir1/experiment_designs.pdf)

Tung, T'ung-Ho. 1957. Phonology of the Amoy dialect (Xiamen fangyan de yinyun), *Bulletin of the Institute of History and Philology. Academia Sinica* 29.1:231-253.

Tung, T'ung-Ho. 1959. Four South Min dialects (Si ge Minnan Fangyan). *Bulletin of the Institute of History and Philology, Academia Sinica* 30:729-1042.

Wang, Hsu Samuel, and Hui-Chuan Liu. 2010. The morphologization of liaison consonants in Taiwan Min and Taiwan Hakka. *Language and Linguistics* 11.1:1-20.

Wang, Hsu Samuel, and Shih-Ting Kao. 2004. On the phonological status of geminated consonants in Taiwan Min. In Y. C. Lin et al. (eds.), *Studies on Sino-Tibetan Languages: Papers in Honor of Professor Hwang-cherng Gong on His Seventieth Birthday*, pp.821-834. Taipei: Institute of Linguistics, Academia Sinica.

Wang, Shih-Ping. 1991. Tone segment interaction: Phonetics and phonological aspects of gemination in Taiwanese, Paper presented at *NACCL 3*, Ithaca, Cornell University.

Wang, Shih-Ping. 1995. Tone-segment interaction: Notes on simplification, In Tsao, F. F., Tsai, M. H. (Eds.), *First International Symposium on Languages in Taiwan*. pp.487-512. Taipei: The Crane Publishing.

Yang, Hsiu-Fang. 1991. *Manuscripts of the Grammar of Southern Min in Taiwan*. Taipei: Da-an Press.

Yip, Moira. 1980. *The Tonal Phonology of Chinese*. MIT. Doctoral dissertation.

Zhang, Zhen-Xing. 1993. *The Southern Min Dialects in Taiwan*. Taipei: Mansbook Press.

[Received 02 December 2014; revised 20 April 2015; accepted 26 June 2015]

*Lin Pingyou*
*Department of English Studies*
*Mingdao University*
*Changhua, Taiwan*
*lin6607@mdu.edu.tw*

## APPENDIX A

### F1 Frequencies of [kim$^{\underline{ma}}$a] and [$\underline{m^{ma}a}$]

F1 frequencies at different time points in 'gold' [kim$^{\underline{ma}}$a].

| Time(s) | Frequency(Hz) | Frequency |
|---|---|---|
| 0.0272 | 983.0435 | N/A |
| 0.0335 | 262.5887 | -720.4548 |
| 0.0397 | 272.7314 | 10.1427 |
| 0.0460 | 277.0938 | 4.3624 |
| 0.0522 | 273.1896 | -3.9041 |
| 0.0585 | 278.5936 | 5.4040 |
| 0.0647 | 286.9658 | 8.3722 |
| 0.0710 | 284.9715 | -1.9943 |
| 0.0772 | 285.8568 | 0.8853 |
| 0.0835 | 294.7401 | 8.8833 |
| 0.0897 | 321.7143 | 26.9742 |
| 0.0960 | 341.1037 | 19.3894 |
| 0.1022 | 326.7396 | -14.3642 |
| 0.1085 | 313.7616 | -12.9780 |
| 0.1147 | 256.8235 | -56.9380 |
| 0.1210 | 225.4825 | -31.3410 |
| 0.1272 | 225.6335 | 0.1510 |
| 0.1335 | 225.4014 | -0.2322 |
| 0.1397 | 233.1910 | 7.7896 |
| 0.1460 | 239.0034 | 5.8124 |
| 0.1522 | 240.5172 | 1.5139 |
| 0.1585 | 237.7518 | -2.7655 |
| 0.1647 | 237.7937 | 0.0420 |
| 0.1710 | 239.0636 | 1.2699 |
| 0.1772 | 234.2187 | -4.8449 |
| 0.1835 | 235.8215 | 1.6028 |
| 0.1897 | 246.5159 | 10.6943 |
| 0.1960 | 331.8794 | 85.3635 |
| 0.2022 | 493.2503 | 161.3709 |
| 0.2085 | 587.2845 | 94.0342 |
| 0.2147 | 634.8721 | 47.5876 |
| 0.2210 | 660.9060 | 26.0339 |
| 0.2272 | 687.6811 | 26.7751 |
| 0.2335 | 702.3803 | 14.6992 |
| 0.2397 | 705.6240 | 3.2437 |
| 0.2460 | 709.4479 | 3.8239 |
| 0.2522 | 704.2583 | -5.1897 |
| 0.2585 | 701.0688 | -3.1895 |
| 0.2647 | 694.7938 | -6.2750 |
| 0.2710 | 690.9735 | -3.8203 |
| 0.2772 | 682.1399 | -8.8337 |
| 0.2835 | 683.2499 | 1.1101 |
| 0.2897 | 673.6731 | -9.5769 |
| 0.2960 | 669.2948 | -4.3783 |
| 0.3022 | 671.9643 | 2.6695 |
| 0.3085 | 650.5716 | -21.3926 |
| 0.3147 | 640.1968 | -10.3749 |
| 0.3210 | 642.0741 | 1.8774 |
| 0.3272 | 617.6013 | -24.4728 |
| 0.3335 | 601.3875 | -16.2138 |
| 0.3397 | 637.3489 | 35.9614 |
| 0.3460 | 641.8186 | 4.4698 |
| 0.3522 | 663.3178 | 21.4991 |

F1 frequencies at different time points in 'horse' [m$^{ma}$a].

| Time(s) | Frequency(Hz) | Frequency |
|---|---|---|
| 0.0269 | 301.4294 | N/A |
| 0.0332 | 293.6989 | -7.7305 |
| 0.0394 | 178.3076 | -115.3913 |
| 0.0457 | 357.9552 | 179.6476 |
| 0.0519 | 659.1017 | 301.1465 |
| 0.0582 | 754.3000 | 95.1983 |
| 0.0644 | 779.0859 | 24.7859 |
| 0.0707 | 782.8953 | 3.8094 |
| 0.0769 | 795.1107 | 12.2154 |
| 0.0832 | 793.3320 | -1.7787 |
| 0.0894 | 790.0828 | -3.2492 |
| 0.0957 | 788.5040 | -1.5787 |
| 0.1019 | 789.3761 | 0.8720 |
| 0.1082 | 787.8948 | -1.4813 |
| 0.1144 | 806.0274 | 18.1326 |
| 0.1207 | 827.3628 | 21.3354 |
| 0.1269 | 836.6147 | 9.2519 |
| 0.1332 | 803.7056 | -32.9091 |
| 0.1394 | 826.8109 | 23.1053 |
| 0.1457 | 847.1487 | 20.3378 |
| 0.1519 | 809.2611 | -37.8876 |
| 0.1582 | 729.7478 | -79.5133 |
| 0.1644 | 759.9569 | 30.2090 |
| 0.1707 | 763.0554 | 3.0986 |
| 0.1769 | 772.5603 | 9.5049 |
| 0.1832 | 733.3155 | -39.2448 |
| 0.1894 | 678.6158 | -54.6997 |
| 0.1957 | 622.1577 | -56.4581 |
| 0.2019 | 591.1454 | -31.0124 |
| 0.2082 | 449.9023 | -141.2431 |

## APPENDIX B
## ABX Discrimination Questionnaire

**Instruction:** The purpose of this experiment is to investigate whether the target sound X is similar to the A token or the B token. It is not a quiz, so relax. Each question has two boxes ☐A and ☐B, so please mark your answer ■ on either A or B. Each question will be broadcast twice, so listen carefully before marking your choice. Thank you very much for your cooperation.

**Personal Information**

Gender: ☐Male ☐Female

Major: _____ Year: _____

Age: _____

Taiwanese Proficiency: ☐Very Good ☐Good ☐Fair ☐Poor ☐Very Poor

| No. | X=A | X=B | No. | X=A | X=B |
|---|---|---|---|---|---|
| 1 | ☐A | ☐B | 25 | ☐A | ☐B |
| 2 | ☐A | ☐B | 26 | ☐A | ☐B |
| 3 | ☐A | ☐B | 27 | ☐A | ☐B |
| 4 | ☐A | ☐B | 28 | ☐A | ☐B |
| 5 | ☐A | ☐B | 29 | ☐A | ☐B |
| 6 | ☐A | ☐B | 30 | ☐A | ☐B |
| 7 | ☐A | ☐B | 31 | ☐A | ☐B |
| 8 | ☐A | ☐B | 32 | ☐A | ☐B |
| 9 | ☐A | ☐B | 33 | ☐A | ☐B |
| 10 | ☐A | ☐B | 34 | ☐A | ☐B |
| 11 | ☐A | ☐B | 35 | ☐A | ☐B |
| 12 | ☐A | ☐B | 36 | ☐A | ☐B |
| 13 | ☐A | ☐B | 37 | ☐A | ☐B |
| 14 | ☐A | ☐B | 38 | ☐A | ☐B |
| 15 | ☐A | ☐B | 39 | ☐A | ☐B |
| 16 | ☐A | ☐B | 40 | ☐A | ☐B |
| 17 | ☐A | ☐B | 41 | ☐A | ☐B |
| 18 | ☐A | ☐B | 42 | ☐A | ☐B |
| 19 | ☐A | ☐B | 43 | ☐A | ☐B |
| 20 | ☐A | ☐B | 44 | ☐A | ☐B |
| 21 | ☐A | ☐B | 45 | ☐A | ☐B |
| 22 | ☐A | ☐B | 46 | ☐A | ☐B |
| 23 | ☐A | ☐B | 47 | ☐A | ☐B |
| 24 | ☐A | ☐B | 48 | ☐A | ☐B |

## APPENDIX C

### Mean Accuracy of the Two Target Words

| Listener | kim a | kim ma |
|---|---|---|
| #1 | 0.83 | 0.83 |
| #2 | 0.54 | 0.71 |
| #3 | 0.79 | 0.63 |
| #4 | 0.46 | 0.33 |
| #5 | 0.42 | 0.38 |
| #6 | 0.92 | 0.75 |
| #7 | 0.83 | 0.83 |
| #8 | 0.67 | 0.58 |
| #9 | 0.83 | 0.67 |
| #10 | 0.88 | 0.83 |
| #11 | 0.79 | 0.63 |
| #12 | 0.79 | 0.92 |
| #13 | 0.79 | 0.67 |
| #14 | 0.83 | 0.75 |
| #15 | 0.79 | 0.88 |
| #16 | 0.79 | 0.75 |
| #17 | 0.75 | 0.88 |
| #18 | 0.42 | 0.33 |
| #19 | 0.71 | 0.79 |
| #20 | 0.75 | 0.75 |
| #21 | 0.88 | 0.96 |
| #22 | 0.79 | 0.75 |
| #23 | 0.83 | 0.88 |
| #24 | 0.83 | 0.75 |
| #25 | 0.92 | 0.75 |
| #26 | 0.83 | 0.67 |
| #27 | 0.92 | 0.88 |
| #28 | 0.79 | 0.71 |
| #29 | 0.83 | 0.96 |
| #30 | 0.54 | 0.42 |
| #31 | 0.83 | 0.83 |
| #32 | 0.67 | 0.46 |
| #33 | 0.83 | 0.75 |
| #34 | 0.88 | 0.71 |
| #35 | 0.96 | 0.88 |
| #36 | 0.79 | 0.79 |
| #37 | 0.96 | 0.96 |
| #38 | 0.75 | 0.71 |
| #39 | 0.88 | 0.67 |
| #40 | 0.83 | 0.63 |
| #41 | 0.96 | 0.92 |
| #42 | 0.79 | 0.83 |
| #43 | 0.96 | 0.71 |
| #44 | 0.67 | 0.50 |
| #45 | 0.79 | 0.88 |
| #46 | 0.83 | 0.92 |
| #47 | 0.46 | 0.54 |
| #48 | 0.88 | 0.75 |
| #49 | 0.96 | 0.79 |
| #50 | 0.79 | 0.83 |
| #51 | 0.92 | 0.75 |
| #52 | 0.67 | 0.75 |
| #53 | 0.50 | 0.54 |
| #54 | 0.50 | 0.54 |
| #55 | 0.83 | 0.88 |
| #56 | 0.88 | 0.88 |
| #57 | 0.50 | 0.54 |
| #58 | 0.75 | 0.83 |
| #59 | 0.96 | 0.92 |
| AVG | 0.77 | 0.73 |

# 三個臺語詞「金鴨」、「金仔」、「金馬」的語音表徵分析

林秉宥
明道大學

本文以 Goldsmith (1976) 所提出的自主音段音韻學架構分析三個臺語詞彙「金鴨」、「金仔」、「金馬」的語音表徵形式。由語音聲學、發音、及聽覺感知分析結果顯示，三詞彙 [m] 時長在統計上呈現顯著性差異。然而，莫拉理論在應用上卻有其侷限性；它或許可解釋日語詞 *kita* [ki(μ)][ta(μ)] 與 *kitta* [kit(μ)(μ)][ta(μ)] 的時長差異，卻無法解釋臺語詞「金仔」[kim(μμ)ma(μ)] 與「金馬」[kim(μμ)][ma(μ)] 的時長差異。因此本文乃提出連音「二空缺」及「雙向同化規則」來解釋此時長差異，即：時長差異乃導因於不同韻律單位的語音轉折，因此三詞彙的 [m] 時長差異即為「金鴨」[kim][ˀaˀ] vs.「金仔」[kimᵐᵃa] vs.「金馬」[kim]ᵐᵐ[ma]。

關鍵詞：臺語、語音表徵、自主音段架構、連音、雙向同化規則、莫拉理論