The Transition Sound of the Taiwanese Suffixation Word *kim.a* 'gold'*

Ping-you Lin

MingDao University

The purpose of this research is to investigate the phonetic representation of Taiwanese words triggered by the suffix *a*, using *kim.a* 'gold' as a case study. This paper proposes a new perspective on the autosegmental framework (Goldsmith 1976), and this new model emphasizes the role of a transition sound between the two morphemes (Lin 2006). The proposed account claims that the boundary between these morphemes is occupied by a transition sound (here called "two-slots") through bidirectional spreading as [kim^{ma}a]. The "two-slots" hypothesis is based on the acoustic evidence; most importantly, it solves the phonetic distinction between the two words [kim^{ma}a] 'gold' and [kimma] 'place names'.

Key words: affixal phonology, autosegmental framework, two-slots hypothesis, transition sound, bidirectional spreading

* This study is a revised version of part of my dissertation (Lin 2006). First of all, I would like to express my deepest gratitude to my father Na, Tsĩ-tsng (林金村) for his academic guidance and everlasting support. Next, I am grateful to the anonymous reviewers for their significant comments and suggestions. Finally, I am really thankful to Stephen Bobroff and Caroline Deveau who helped with editing and offered comments as well. All remaining errors are, of course, mine.

1. Introduction

The Taiwanese *a* suffixation[1] or affixal phonology has been extensively studied in the last two decades (Lin 2004), and Goldsmith's Autosegmental Phonology (1976) is often used as an analysis model. The autosegmental framework, which is also known as non-linear phonology, treats phonological representations as a multi-dimensional rather than a linear display of feature matrices. More precisely, some phonological features that are regarded as independent segments or autosegments can be taken away from their original segments and arranged on separate autonomous tiers, so that rules applying to one tier may not necessarily apply to another. These autosegments are linked to each other by association lines to indicate interrelationship among them. However, tones on tonal tier, nasals on nasality tier, or other prosodic features on other tiers are not covered in this paper. Besides, this paper only focuses on the segmental behavior of two segments across morpheme boundaries rather than two segments within a syllable.

After the suffix *a* is attached to a stem, a "linker-like brief sound" will appear in the morpheme boundary. In other words, the linker-like sound is the "transition sound" that appears between the two morphemes, connecting the final sound of the stem to the following suffix *a*, so the whole word can be articulated in a continuous manner. If *kim.a* '金仔gold'[2] is used as an example word, the description of this linking sound is always described as an additional segment of [m] occurring before the suffix, such as [ki ma], [kimma], or [kima]. These are based on the established approaches of phonological representations stimulated by the suffix *a*. The three approaches are known as resyllabification (Yip 1980), gemination (Ang 1985; Cheng 1997; Cheng and Cheng Xie 1977; Chiang 1992; Chung 1996; Lin 1989; Yang 1991; Zhang 1993) and ambisyllabicity (Kao 2003; Tung 1957, 1959; Wang 1991, 1995).

Using Clements & Keyser's (1983) CV phonology, the three approaches are presented in a three-tiered autosegmental structure in diagram (1). First, the resyllabification in (1a) is a segmental process in which the coda [m] is removed from the first syllable and connected to the following syllable; the coda that the suffix receives as its onset is an additional sound before the suffix. Second, the gemination in (1b) indicates a process whereby the final segment [m] occupies two successive timing slots on the skeletal tier. Thus an extra segment appears before the suffix to form the syllable [ma]. Finally, the ambisyllabicity in (1c) shows a phonological process in

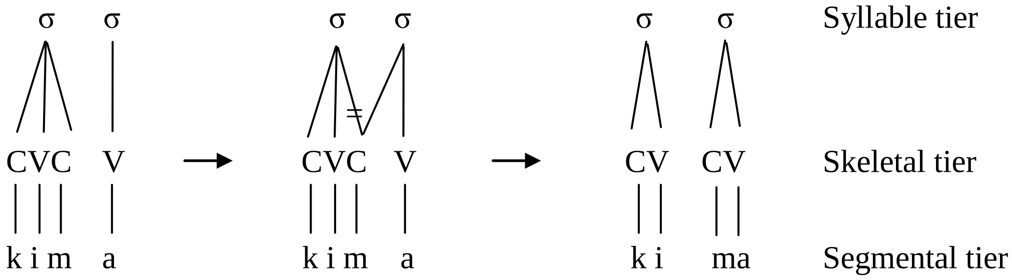
[1] The Taiwanese diminutive suffix *a* denotes the meaning of smallness or is merely a generic term. The tones in the data are omitted since they are irrelevant to the topic here.

[2] The dot before the suffix *a* is adopted here to signify the suffixation marker.

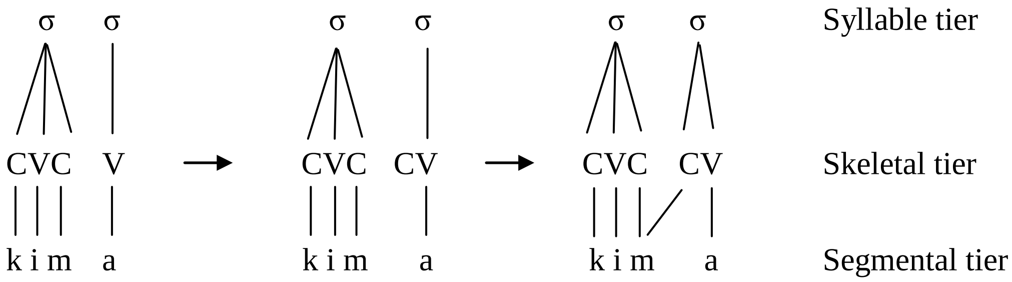
which the final segment [m] is shared simultaneously by two contiguous syllables, so that the suffix acquires the coda as its onset by means of the linkage.

(1)Three-tiered autosegmental representations of the three approaches

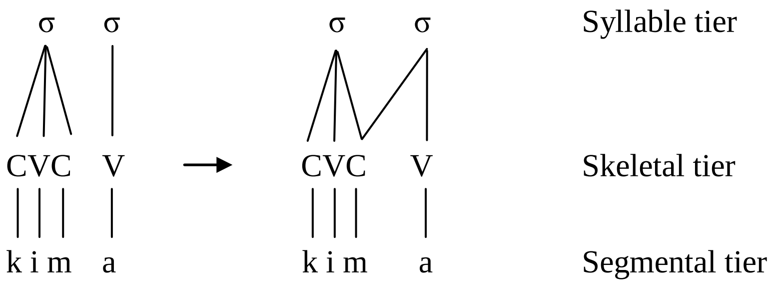
a. Resyllabification



b. Gemination



c. Ambisyllabicity



As the vowel initial suffix *a* attaches to a stem, it triggers a linking sound between the two morphemes. Each of these three analyses suggests a possible interpretation of what a native speaker perceives the linking sound to be. Nevertheless, they may fail to account for the data satisfactorily. First, the resyllabification approach delinks the coda [m] from the first syllable and reconnects it with the second syllable and derives [ki ma]; however, this description does not reflect the linguistic knowledge of native speakers because the word is articulated as /kim/ '金' plus /a/ '仔' by the native speakers, not /ki/ '枝' plus /ma/ '馬'. Further, its autosegmental representation in (1a) yields a juncture between the first syllable [ki] and the second syllable [ma] since it lacks an association line to link the two syllables. Any *a*-attached words are articulated in continuous speech, therefore, the *a*-attached word *kim.a* is pronounced in continuous speech between [kim] and [a] rather than a juncture (pause) appearing in the morpheme boundary between [ki] and [ma].

Second, the gemination analysis claims that the linking sound is perceived as a C-slot insertion spread from the coda [m] to fulfill the preferred conventional CV syllable template as [kimma]. Yet, the problem is: if the phonological form of the word is represented as an extra C on the skeletal tier, how would one differentiate this "true gemination" from the fake gemination of the word *kim-ma*[3] '金馬 the abbreviation of Kimmng and Matsoo[4]' since the derived outputs of both have two C-slots on the skeletal tier, i.e., [ki<u>mm</u>a] 'gold' vs. [ki<u>mm</u>a] or [ki<u>m</u>][<u>m</u>a] 'place names'? Actually, the relative duration of a fake geminate (i.e., *kim-ma*) is significantly longer than a true geminate (i.e., *kim.a*), which is about 1.38:1 (Chiang 1992:79) and any phonological rule that derives the phonetic output should indicate such difference. However, the gemination analysis cannot efficiently reflect this difference (cf. the duration of [m] in Figure 1 and Figure 2).

[3] The hyphen used here indicates the compound word marker.
[4] The Taiwanese Romanization System is used in this study.

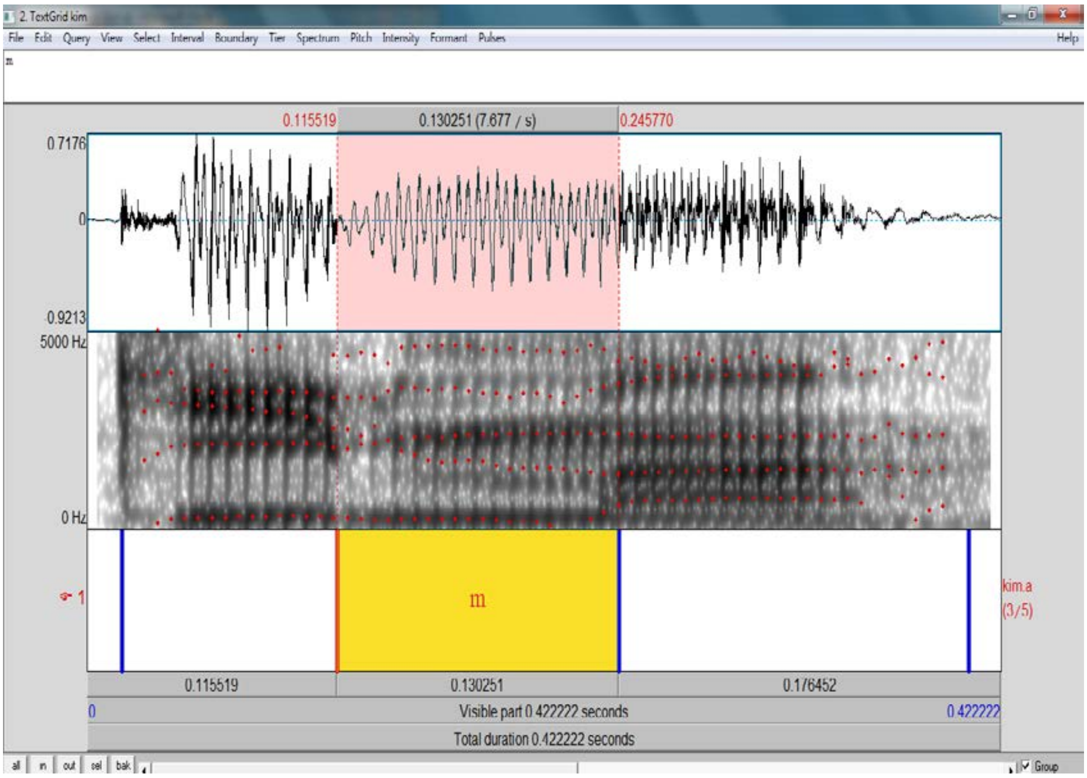


Figure 1. The [m] duration of 0.1303s for the word *kim.a*.

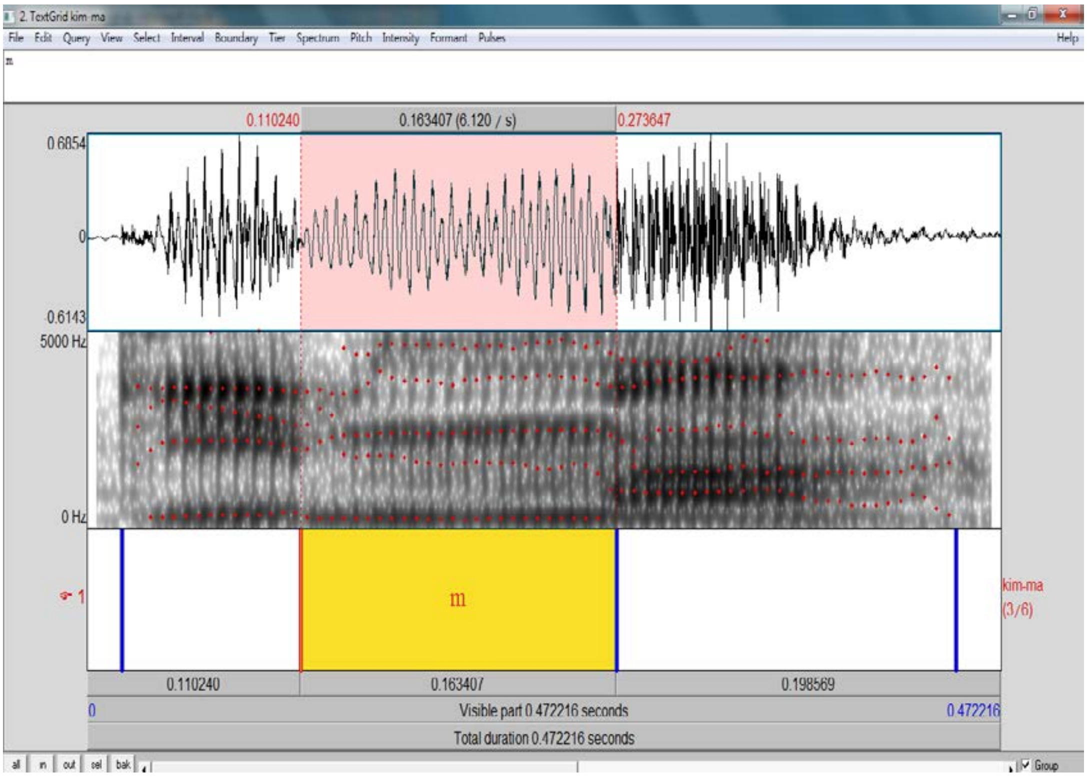


Figure 2. The [m] duration of 0.1634s for the word *kim-ma*.

Except for the similar phonetic realizations of the two words, the additional C-slot claimed from the gemination analysis lacks acoustic evidence. In other words, how can we identify from the spectrogram that the C-to-V linking sound is a consonant? Is the formant transition from C-to-V a consonant sound? These questions still remain unsettled.

Third, as to the ambisyllabicity approach, the segment [m] is argued to belong simultaneously to the coda of the first syllable and the onset of the second syllable. Though this description captures the fact that the stem and the suffix indeed are linked in some way, the phonological representation on the skeletal tier [kima] may lack acoustic evidence, too. In other words, how can the coda and onset [m] be proved acoustically?

In English, perhaps there is a somewhat parallel comparison. Kahn (1976) exemplifies the structure of the phrase *hock it*, as the following diagram illustrates:

(2) hock it

```
h a k   i t
 \|/     \/
  S      S        (Kahn 1976: 30)
```

In (2), the [k] is not linked to the succeeding syllable. However, in ordinary speech the sound [k] is regarded to be associated with the syllable *it*. Kahn states that in inter-word syllabic linking, the final /p, t, k/ in ordinary speech must associate with a following word-initial vowel. He clarifies:

> However, whether or not Rule V applies to make /p, t, k/ syllable-initial, its original syllable-final structure must remain, excluding it from the domain of the aspiration rule. This is also a correct prediction, for whether *tap Ann*, *hit Ann*, *shake Ann* are pronounced with or without pause, /p, t, k/ must not be aspirated. (p. 53)

Three conclusions can be drawn from Kahn's statement. First, the segmental behavior of *hock it* is not resyllabification since its original syllable-final /k/ must remain, thus, it can be excluded from the domain of the aspiration rule. Second, the segmental behavior of *hock it* is not gemination because in inter-word syllabic linking, the final /k/ in ordinary speech must associate with a following word-initial vowel rather than generate an onset [k] to be the aspirated sound [k^h]. Third, the segmental behavior of *hock it* may be ambisyllabicity. The problem is that the ambisyllabic /k/ has to link with the following word-initial vowel /i/ to fulfill the maximal onset principle; at the same time, /k/ also has to fulfill Kahn’s statement, “…whether *tap Ann*, *hit Ann*, *shake Ann* are pronounced with or without pause, /p, t, k/ must not be aspirated.” Based on this statement, how can /k/ be linked with the following syllable without being aspirated since /k/ is also the onset of the following syllable? Analogously, in the Taiwanese word *kim.a* the syllable-final structure [m] in *kim* must remain and be associated with the syllable *a*; in the meantime, the [m] has to be prevented from combining with the succeeding syllable to form the new syllable *ma* because Kahn's description indirectly proves that the new syllable [k^{h}ɪt] cannot be derived. Though the ambisyllabic /k/ may be acceptable, the sound may still require further acoustic evidence.

Regarding the problems listed above, this paper proposes the “two-slots hypothesis,” which emphasizes the two brief segments occupying the morpheme boundary. In other words, the linking sound equals the two timing slots. In this study, I argue that the coda must remain; nevertheless, the coda has to avoid joining with the suffix to form a new syllable. The solution to it perhaps can be described in the following phonological process: After *a* suffixes the stem *kim*, it triggers both the final segment [m] and the suffix [a] to occupy the two timing slots ○○ in the morpheme boundary as the linking sound. So the derived segment from [m] is the first slot, and the derived segment from [a] is the second slot, and the two derived segments can unify as the C-to-V linking sound [^{ma}]. Specifically, the first syllable /kim/, on the one hand, can link with the second syllable [a] through the transition sound [^{ma}]; on the other hand, the second slot [^{a}] in the transition sound can prevent the first slot [^{m}] from permeating the second syllable to form the new syllable [ma]. As a result, the surface form of [kim^{ma}a] ‘gold’ can be efficiently differentiated from that of [kimma] ‘place names.’ The “two-slots” proposal will further be clarified in the Results and Discussion section.

2. Methods

This study essentially adopts acoustic analysis to provide another possible interpretation of the same word examined in earlier studies. The F1 transition and the wave form of the word will provide evidence as to the validity of the transition sound in the two-slots hypothesis. This may offer an adequate description of what the phonetic realization of the “stem + suffix” formation is.

Twelve native speakers were chosen as the participants for this study. The number of participants was based on Ladefoged’s (2003: 14) suggestion, “… half a dozen speakers of each sex” will be enough because “there may be systematic differences between male and female speech". The age of the participants ranged from 20's to 70's; the predominant languages used in their daily lives were both Taiwanese and Mandarin. The background of the twelve participants is presented in Table 1.

Table 1. Profile of the twelve participants

	Par #1	Par #2	Par #3	Par #4	Par #5	Par #6	Par #7	Par #8	Par #9	Par #10	Par #11	Par #12
Sex	M	F	M	F	M	F	M	F	M	F	M	F
Age	52	46	37	61	68	24	39	44	74	69	63	69
Lan	T/M	T/M	T/M	T/M	T/M	T/M	T/M	T/M	T	T/M	T/M	T/M

Note. Lan: Language. T: Taiwanese. M: Mandarin.

The recordings were conducted at MingDao University's radio studio, which had a relaxing and quiet atmosphere. The participants' task was to identify the object of gold in a photograph as well as other objects presented in photographs, and name each of them (In this paper, only the word *kim.a* is discussed). Each object was represented in two different photographs in case one of them could not be recognized. All of the photographs were randomly presented through a PowerPoint slideshow. After an expected word was produced by each participant, the researcher showed the next photograph. The participants were asked to say the names of the objects into a microphone (Audio Technica AT4040) and their voices were recorded with a digital recorder (Marantz CDR 300). The sampling rate used here is 44100Hz. They were asked to produce the names in a voice which was clear and loud enough for the acoustic analysis. The photographs were shown one after another until all the photographs were identified, and the audio recorder was turned off. The time for data collection was five minutes for each participant.

Twelve voice samples for the word *kim.a* were collected and they were edited by Praat software (http://www.praat.org, version 5.3.23). The criteria used to segment the transition from the preceding nasal and the following vowel is based on the F1 (first formant) transition, in which the frequency increases dramatically from lower figures of the [m] to higher figures of the vowel [a]. If the frequency difference of F1 rises over 20Hz, it would be regarded as where the transition sound occurs in this study. In addition to the formant shift, the comparison of the two different types of wave forms for the transition sound indicates the sound is composed of two slots as well. As a result, the first formant and the sound wave of the data can be observed and analyzed.

3. Results and Discussion

The "two-slots hypothesis" can be tested according to the acoustic analysis of the word. The spectrogram of the word *kim.a* is demonstrated as below:

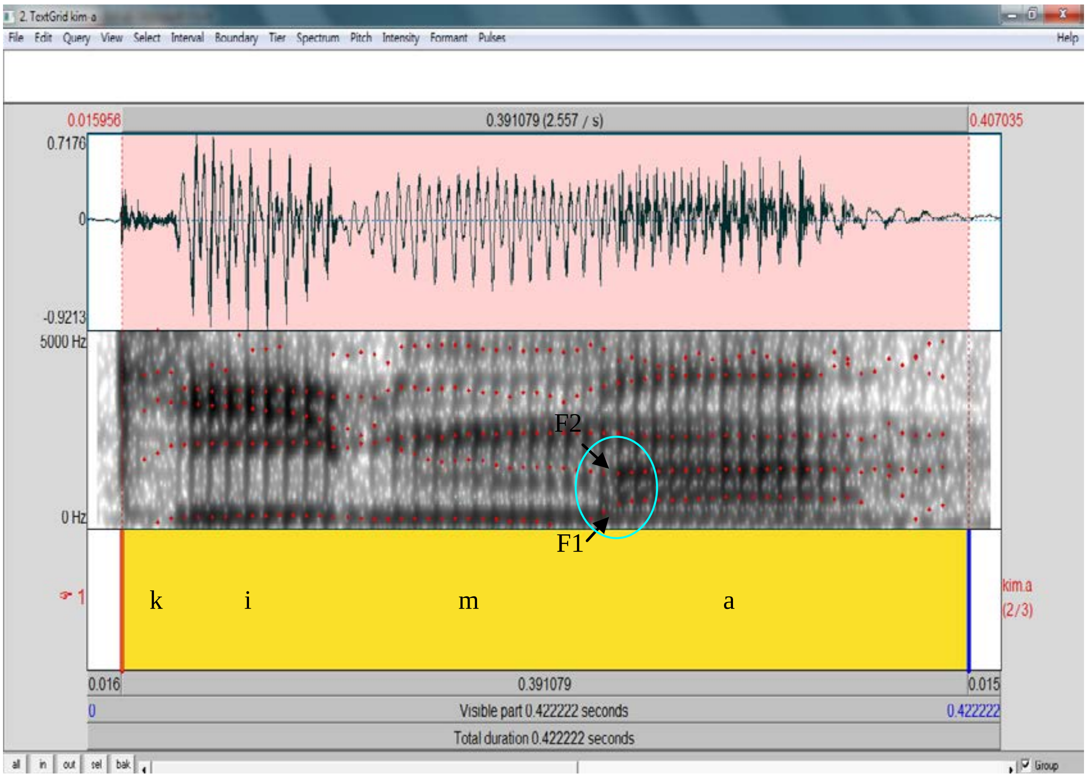


Figure 3. The spectrogram of the word *kim.a* articulated by participant #1.

In Figure 3, the first formant (F1) frequency and the second formant (F2) frequency shift remarkably within the blue circle ; however, the tongue height in the blue circle (related with the F1) (Ladefoged 2006; Pickett 1999; Stevens 1999; among others) changes more evidently than the tongue front (related with the F2). Specifically, the frequency changes of F1, for example, from 0.2267s[5] (253.77Hz) to 0.2705s (797.34Hz) is 543.57Hz; whereas the frequency changes of F2 from 0.2267s (1463.27Hz) to 0.2705s (1444.58Hz) is 18.69Hz. As a result, only the F1 is used for segmentation in this paper. The sound that transits from the higher tongue position of [m] to the lower tongue position of [a] in the blue circle belongs to neither the first syllable *kim* nor the second syllable *a*. It denotes the position where the transition sound happens. The formant transition indicates that the formation of the word can further be divided into three parts and the three parts are characterized in Figure 4.

5 For unit seconds, four decimal places are rounded, while two decimal places are rounded for the Hz values.

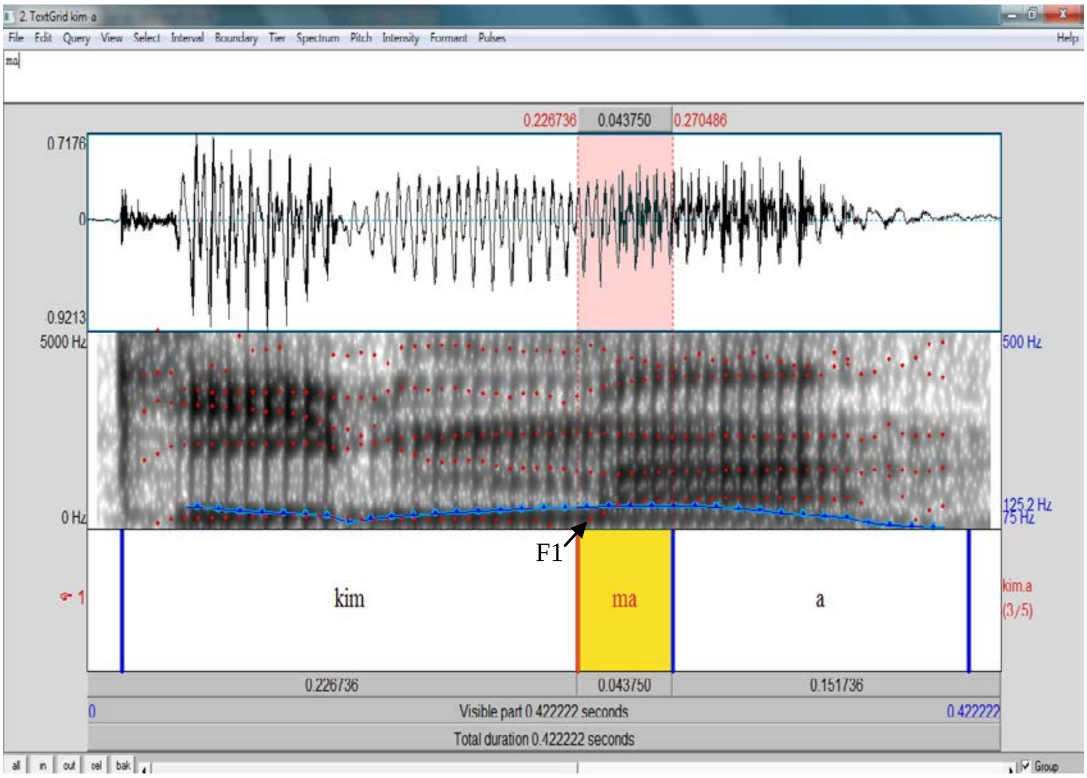


Figure 4. The wave form, formant (the red dots), pitch (the blue horizontal line), and spectrum of the word *kim.a* articulated by participant #1.

In Figure 4, the F1 rises sharply from 0.2267s to 0.2705s. Such formant transition indicates that the tongue changes from one steady position to another steady position. To pronounce the word in a continuous way, the lips are closed for the coda [m] and its F1 is steady before 0.2267s. After 0.2267s, the lips are apart and the tongue moves downward, so the F1 starts to rise. As the tongue moves lower and lower, the formant frequency increases higher and higher. Finally, the oral cavity expands almost to its biggest space which indicates that the tongue moves nearly to its lowest position at 0.2705s; the suffix [a] begins and the formant becomes steady again. Consequently, the F1 transiting from 0.2267s to 0.2705s is the transition sound [^{ma}], and the sound is significant because its F1 differs dramatically from that of both the coda [m] and the suffix [a].

Since F1 transition correlates with the tongue height, the shift of F1 therefore can be utilized to delimit the borders of the transition sound [^{ma}]. The F1 frequencies of the transition sound are shown below:

Table 2. The noticeable figures of frequency difference

Time_s	F1_Hz	Frequency difference
0.0160	N/A	N/A
…	…	…
0.2017	241.76	3.43
0.2080	244.64	2.89
0.2142	247.30	2.65
0.2205	252.57	5.28
0.2267	253.77	1.20
0.2330	287.74	33.97
0.2392	371.06	83.32
0.2455	596.41	225.34
0.2517	650.15	53.74
0.2580	691.22	41.07
0.2642	751.16	59.94
0.2705	797.34	46.17
0.2767	811.72	14.39
0.2830	828.32	16.60
0.2892	834.59	6.24
0.2955	803.25	-31.30
0.3017	804.93	1.68
…	…	…
0.4070	N/A	N/A

In Table 2, some noticeable figures of frequency difference, including 33.97Hz, 83.32Hz, 225.34Hz, 53.74Hz, 41.07Hz, 59.94Hz, and 46.17Hz, are remarkable (Note: As the definition in Methods, if the frequency difference of F1 rises over 20Hz, it would be regarded as where the transition sound occurs). Each of these figures signifies a big change from one time value to the next time value. The first frequency difference of 33.97Hz between 253.77Hz (0.2267s) and 287.74Hz (0.2330s) signifies that 0.2267s is the moment when the lips release and the tongue moves downward, so the F1 starts to rise. The last remarkable value 46.17Hz, which is the frequency difference between 751.16Hz (0.2642s) and 797.34Hz (0.2705s), conveys nearly the lowest tongue position, then the F1 becomes steady after the time value of 0.2705s and the suffix [a] starts. To summarize, these noticeable figures provide the evidence that the F1 transition is located between 0.2267s and 0.2705s, which indicates the transition sound as well. In addition to the F1 transition, the wave form of the transition sound is significant too. The transition sound can be divided into two types of sound waves -- the sound waves of the [m] segment and the sound waves of the [a] segment. This portion of the word is enlarged as follows:

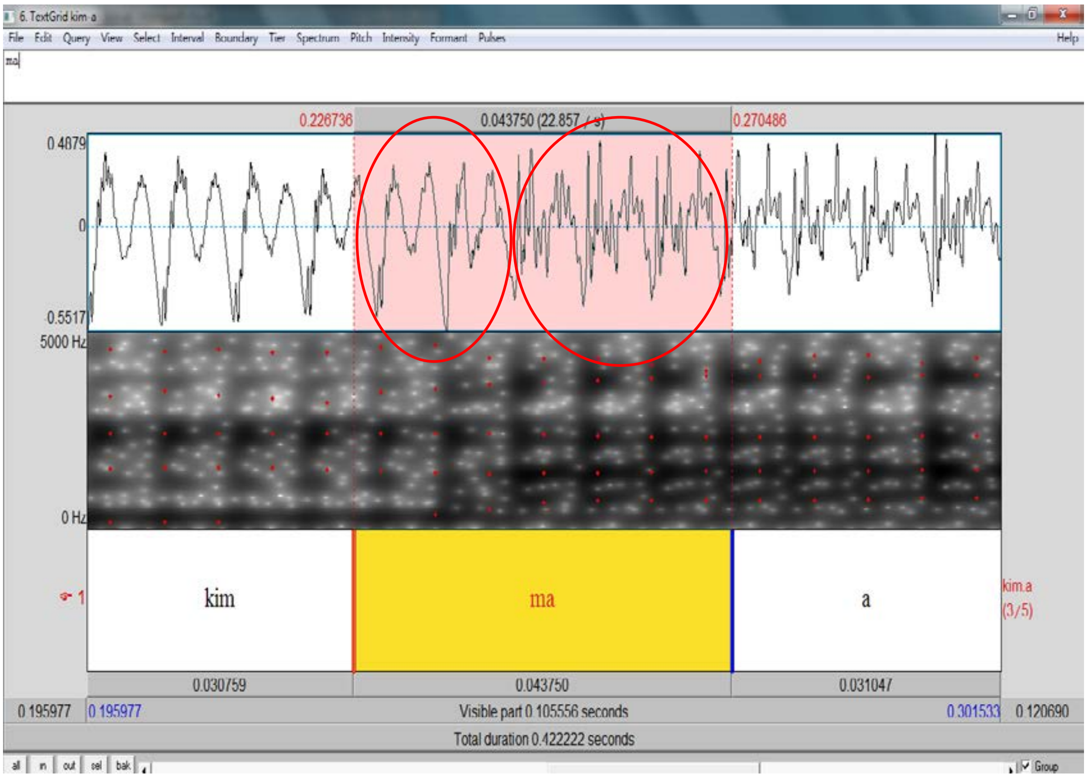


Figure 5. The portion of the enlarged wave form.

As Figure 5 shows, the wave form of the first segment from the transition sound is nearly identical to the wave form of the coda [m]; correspondingly, the wave form of the second segment is roughly similar to the wave form of the suffix [a] as well. The transition sound, therefore, can be regarded as being composed of the two sound segments [^{m}] and [^{a}] -- one resembling the coda [m] and the other resembling the suffix [a]. If the spectrum of the transition sound [^{ma}] is examined, the spectrum of the [^{m}] segment will be less visible than that of the [^{a}] segment. It is because the nasal sound has less energy than the vowel sound. In short, it proves the transition sound is composed of two slots.

Based on Figure 4, Table 2 and Figure 5, three acoustic properties of the transition sound become apparent. First, the transition sound is voiced rather than silent. Second, the duration of the sound is fairly short. Third, the sound is composed of two sound segments. All in all, the three characteristics reveal that the transition sound exists, but that it is short. If the duration of the whole word is measured from the beginning of the

[k] sound (0.0160s) until the end of the sound vibration (0.4070s), the duration for the transition sound is about 11 % of the whole word, which is presented in Table 3:

Table 3. The percentage of each part of the word

	duration	percentage
kim(part A)	0.2108	54%
^{ma}(part B)	0.0438	11%
a(part C)	0.1365	35%
SUM	0.3911	100%

As shown in Table 3, the duration of [kim] (part A), [^{ma}] (part B), and [a] (part C) are approximately 0.2108 seconds (54%), 0.0438 seconds (11%), and 0.1365 seconds (35%), respectively. Although the time span for [^{ma}] (11%) is the shortest among the three parts, it nonetheless exists, and takes up two timing slots on the skeletal tier. If there were no transition sound [^{ma}] in the boundary position, it would sound like the compound word *kim-a* ‘nonexisting word (a word that doesn’t exist)’. The wave form, pitch, and spectrogram records of the compound word are shown in Figure 6. The space between the two blue pitch lines is approximately the boundary position, and it also corresponds to the sound wave, which has less vibration (cf . Figure 4 and Figure 6).

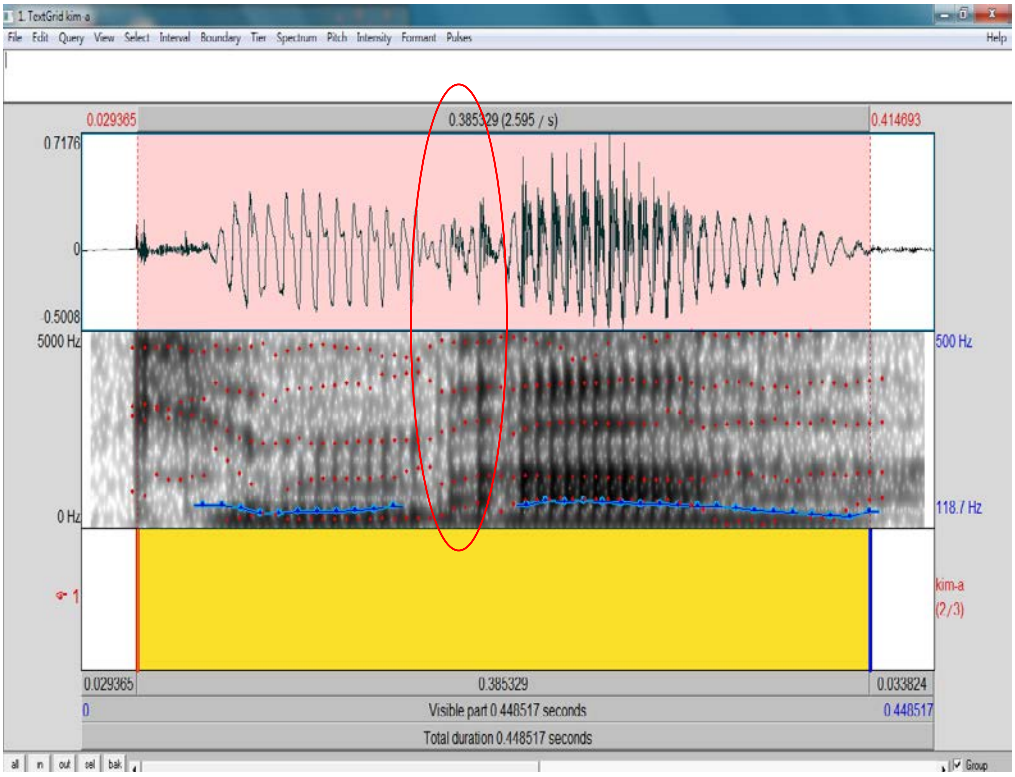
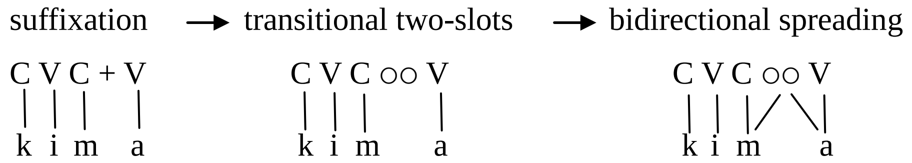


Figure 6. The wave form, pitch, and spectrum of the word *kim-a*.

Next, the transition sound [^{ma}] in part B is produced by bidirectional spreading from the two syllables. This phonological process specifically can be described as follows: the first slot [^{m}] is derived from the consonant [m] in part A and the second slot [^{a}] is derived from the vowel [a] in part C and both segments spread bidirectionally to form the transition sound [^{ma}] that occupies the space in part B.

Since [^{ma}] seems to have the CV structure, we might wonder if it should be considered another separate syllable. Here are the reasons why such a conclusion cannot be drawn: Firstly, since the transition occurs in the morpheme boundary, it should not be treated as a syllable. Secondly, we only consciously pronounce the two syllables [kim] and [a] with a continuous way, and we never intend to pronounce the non-syllable [^{ma}] in between since the transition sound [^{ma}] never appears on the underlying level. Thirdly, [kim] and [a] are considered syllables because their duration tends to be longer. As illustrated, the duration of [^{ma}] (11%) being much shorter than that of its adjacent two syllables (54% and 35%), it is not plausible to consider it a syllable. Since the two slots do not form a regular syllable, they cannot be organized into a prosodic unit like the mora and hence weightless. To sum up, since the transition sound is regarded as a non-syllable, or unintentionally-pronounced unit, the superscript is used to prevent it being misinterpreted. The phonological derivation of the word therefore can be formulated in (3).

(3) The formation of the suffixation word *kim.a*



The discussion so far merely focuses on the performance of participant #1, so the result for the duration of the transition sound may lack reliability. The F1 transition as well as the wave form delimits the transition sound, and the spectrograms for the remaining eleven participants are shown in Appendix. The average of the duration for the overall twelve participants is approximately 10%, as listed in Table 4.

Table 4. The duration of the transition sound articulated by each participant

	Par #1	Par #2	Par #3	Par #4	Par #5	Par #6	Par #7	Par #8	Par #9	Par #10	Par #11	Par #12
T.S.	11%	10%	9%	13%	8%	11%	11%	9%	10%	10%	10%	9%

Note. T.S.: transition sound.

As a result, the phonological process for the target word *kim.a* can be sketched in the following three-tiered autosegmental representation (cf. resyllabification, gemination, and ambisyllabicity in diagram 1):

(4) Two-slots model

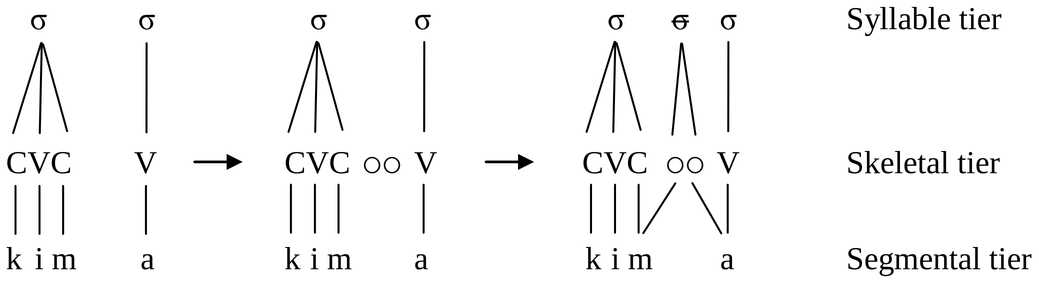


Diagram (4) suggests that the stem [kim] and the suffix [a] are two separate syllables with a boundary between them. If the word is articulated with only the stem *kim* and the suffix *a*, the word will behave as a compound word and a boundary (pause) will occur between the two morphemes, e.g., *kim-a* ‘nonexisting word’. With a boundary between *kim* and *a*, it ends up as the surface form of [kim-   -a], i.e., [kim-boundary-a]. Since this is a suffixation word rather than a compound word, the boundary position between the two morphemes has to be filled with a sound, so that the formation of the “stem + transition sound + suffix” can be pronounced in a continuous fashion. To connect the two morphemes, the non-syllable (shortened as σ̶ in the syllable tier) transition sound has to emerge on the skeletal tier. Therefore, the coda [m] as well as the vowel [a] derive simultaneously to produce the transition sound [^{ma}] to fill the two-slots. The sound which lasts about 10% of the duration fills the boundary position and connects the stem and the suffix; thus the surface form of [kim^{ma}a] eventually surfaces.

4. Conclusion

The existence of the two-slots as the transition sound is evidenced by the acoustic analysis. Both the first formant transition and the diverse wave forms verify the existence of the sound [^{ma}] and this sound is significant since its F1 differs dramatically from that of both the coda [m] and the suffix [a]. Though the duration for this non-syllable sound is short (about 10%), it still exists and takes up two timing slots rather than one timing slot on the skeletal tier. Specifically, the gemination C-slot is actually composed of two sound units and the duration of the C-slot equals the duration of both the [^{m}] plus [^{a}]. Last but not least, the transition sound [^{ma}] solves the phonetic distinction of the two words; namely, *kim.a* and *kim-ma*. Therefore, the existence of the position needs to be observed more precisely.

The two-slots account derived by the bidirectional spreading offers a way to comprehend why there exist different interpretations of resyllabification, gemination, and ambisyllabicity on the same word “kim + a” formation. If observed from the number of timing slots of [m] on the skeletal tier, the answer will be apparent. As demonstrated, there is one timing slot of [m] on the skeletal tier for resyllabification [ki ma] and ambisyllabicity [kima], whereas two timing slots of [m]s on the skeletal tier for gemination [kimma], so the number of [m] always wavers between 1 [m] and 2 [m]s. The results of this study, however, indicate that the number of [m] on the skeletal tier is more than 1 [m] and less than 2 [m]s, as [kim^{ma}a]; i.e., one C-slot pluses one ○-slot roughly result in more than one [m] and less than two [m]s. Perhaps the duration of more than 1 [m] and less than 2 [m]s provides an adequate explanation of why native speakers always hesitate to decide whether it is purely one [m] or two [m]s for the realization of this word (see the experimental discussions of Kao 2003; Wang & Liu 2010).

The theoretical implication of this study is based on the role a transition sound plays in *a*-attached words. In short, between two morphemes, I claim that the boundary can be occupied by either a juncture (a pause) or a transition sound (two-slots). Yet, in the conventional autosegmental framework, little attention has been paid to this position. Accordingly, the proposal can be used to differentiate a compound word from a suffixed word -- juncture vs. transition sound. Specifically speaking, in a word if the stem is followed by an onsetless stem, there must be a juncture instead of a transition sound in between; the word, therefore, is a compound word and uttered in a non-continuous fashion, such as *kim-a* ‘nonexisting word’, *kim-aʔ* ‘golden duck’, *ɔ-aʔ* ‘crow’, *kim-e* ‘golden shoes’, *hɔ-e* ‘rain boots’, *p*^{h}*ɔŋ-i* ‘sofa’, *aŋ-i* ‘witch’, *pue-ɔ* ‘flying fish’, etc. On the other hand, in a word if the stem is followed by an onsetless suffix, the boundary in between must be taken over by a transition sound, so there is no pause between the stem and the suffix; the word, therefore, is a suffixed word and is articulated in a continuous way, such as *kim.a* ‘gold’, *ɔ.a* ‘taro’, *kim.e* ‘gold color’, *hɔ.e* ‘using net for catching fish’, *kan.a* ‘glass bottle’, *paŋ.a* ‘board’, *ab.a* ‘box’, *ts*^{h}*ad.a* ‘thief’, *tig.a* ‘bamboo’, *li.a* ‘plum’, *lu.a* ‘brush’, *e.a* ‘shoe’, *o.a* ‘oyster’, *aʔ.a* ‘duckling’, and so on. Finally, the theoretical implication also implies that the syllable boundaries would be clear-cut and each syllable cannot be permeated by a spreading segment in both the syllable tier and the segmental tier. That is the reason why the rightward spreading from the first syllable to occupy the morpheme boundary and produce a new syllable claimed by the gemination analysis (i.e., [kimma] 金仔=金馬?) is considered inappropriate in this paper.

The pedagogical implication for the non-native speakers can be based on what Na, my father, claims,

> To pronounce the word *kim.a*, what a speaker should intend to articulate is *kim* plus *a* in a continuous way, not *kim* plus *ma*. Non-native students can be guided first to separately practice uttering the stem *kim* and the suffix *a* with a boundary in between, and then in a non-stop manner again pronounce the word *kim.a*. Under this instruction, the transition sound can be naturally and unconsciously derived because we only focus on pronouncing *kim.a* instead of *ki.ma* or *kim.ma*. (Na, Tsĩ-tsng, personal communication, October 15th, 2008).

Moreover, the pedagogical implication can be viewed from Taiwanese teaching textbooks if using Chinese characters with Roman spelling underneath to reflect their pronunciation. An example is shown below:

(5) 金 仔
kim.a

The Chinese and Roman orthographies in (5) assist the abovementioned instruction as well because the Chinese characters stabilize the exact sound of each character. As a result, since 金 is pronounced [kim] and 仔 is pronounced [a] and the [a] should attach to the stem because the dot signifies the suffixation marker, the two characters articulated together can surface the phonetic form of [kim^{ma}a], not [ki ma] nor [kimma]. The expected speech sound, thus, can definitely be produced and can further be differentiated from the word *kim-ma* by students.

For explaining the phonological process of Taiwanese suffixation, the value of using the two-slots approach, by which the transition sound is generated through bidirectional spreading, has been established using *kim.a* as a case study. This already opens a door to a further application of the proposed framework either to other Taiwanese suffixation words with different codas or to other languages like Chinese and English. The affixal phonology will be interesting if compared cross-linguistically. Perhaps the transition sound between two morphemes could be considered a universal phonological phenomenon throughout the human languages.

References

Ang, U. J. 1985. *Studies on the Tones of Taiwan Holo*. Taipei: Zhili Press.

Cheng, L. W. R. 1997. *Taiwanese and Mandarin Structures and Their Developmental Trends in Taiwan I: Taiwanese Phonology and Morphology*. Taipei: Yuanliou Publishing.

Cheng, L. W. R. & Cheng Xie, S. J. 1977. *Phonological Structure and Romanization of Taiwanese Hokkien*. Taipei: Student Book Publishing.

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

Chung, R. F. 1996. *The Segmental Phonology of Southern Min in Taiwan*. Taipei: Crane Publishing.

Clements, G. & Keyser, S. 1983. *CV Phonology: A Generative Theory of the Syllable.* Cambridge, MA: MIT Press.

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

Kahn, D. 1976. *Syllable-based Generalizations in English Phonology*. Doctoral dissertation. MIT.

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

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

Ladefoged, P. 2006. *A Course in Phonetics (5th ed.).* Boston: Thomson Wadsworth.

Lin, P. Y. 2006. *A Study of the Taiwanese Suffix -a*[53]*: An Integrated Autosegmental Analysis*. Doctoral dissertation. De La Salle University.

Lin, Y. H. 1989. *An Autosegmental Treatment of Chinese Segments*. Doctoral dissertation. University of Texas, Austin.

Lin, Y. H. 2004. Chinese affixal phonology: Some analytical and theoretical issues. *Language and Linguistics* 5 (4): 1019-1046.

Pickett, J. M. 1999. *The Acoustics of Speech Communication: Fundamentals, Speech Perception Theory, and Technology*. Boston: Allyn and Bacon.

Stevens, K. 1999. *Acoustic Phonetics*. Cambridge: MIT press.

Tung, T. H. 1957. Xiamen fangyan de yinyun [Phonology of the Amoy dialect]. *Bulletin of the Institute of History and Philology, Academia Sinica* 29 (1): 231-253.

Tung, T. H. 1959. Si ge Minnan fangyan [Four South Min dialects], *Bulletin of the Institute of History and Philology, Academia Sinica* 30: 729-1042.

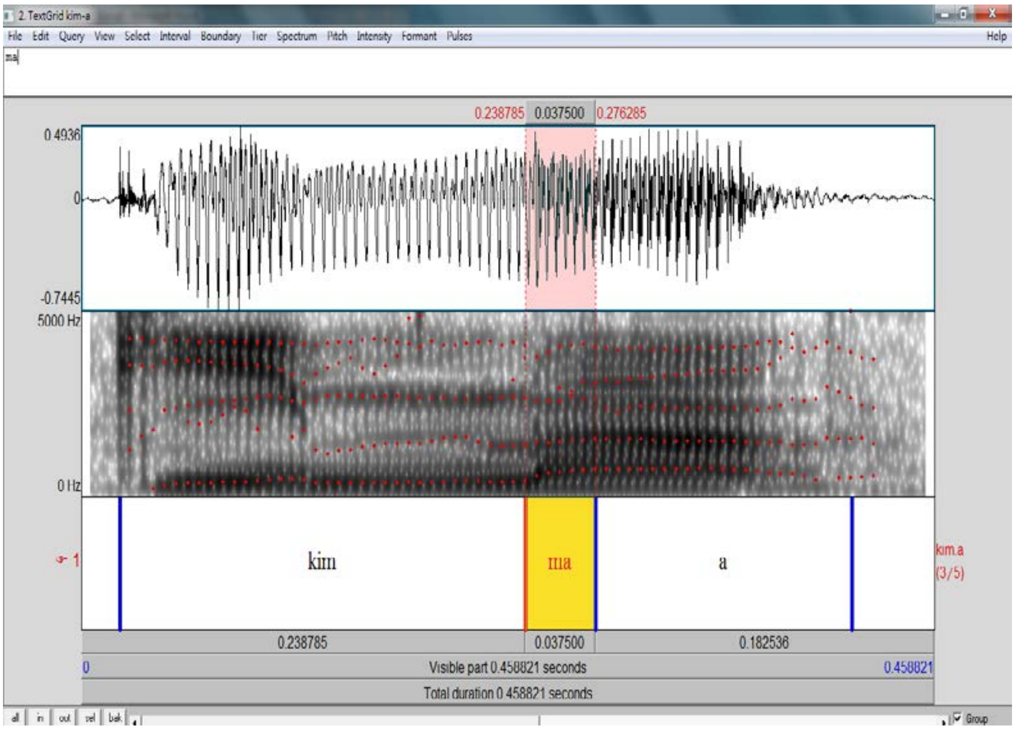
Wang, H. S. & Liu, H. C. J. 2010. The morphologization of liaison consonants in

Taiwan Min and Taiwan Hakka. *Language and Linguistics* 11 (1): 1-20.
Wang, S. P. 1991. Tone segment interaction: Phonetics and phonological aspects of gemination in Taiwanese, Paper presented at *NACCL 3*, Ithaca, Cornell University.
Wang, S. P. 1995. Tone-segment interaction: Notes on simplification, In Tsao, F. F. & Tsai, M. H. (Eds.), *First International Symposium on Languages in Taiwan*. Taipei: Crane Publishing 487-512.
Yang, H. F. 1991. *A Grammar of Taiwan Minnan*. Taipei: Da-an Press.
Yip, M. 1980. *The Tonal Phonology of Chinese*. Doctoral dissertation. MIT.
Zhang, Z. X. 1993. *The Southern Min Dialects in Taiwan*. Taipei: Mansbook Press.

林秉宥
明道大學應用英語學系
lin6607@mdu.edu.tw

Appendix:

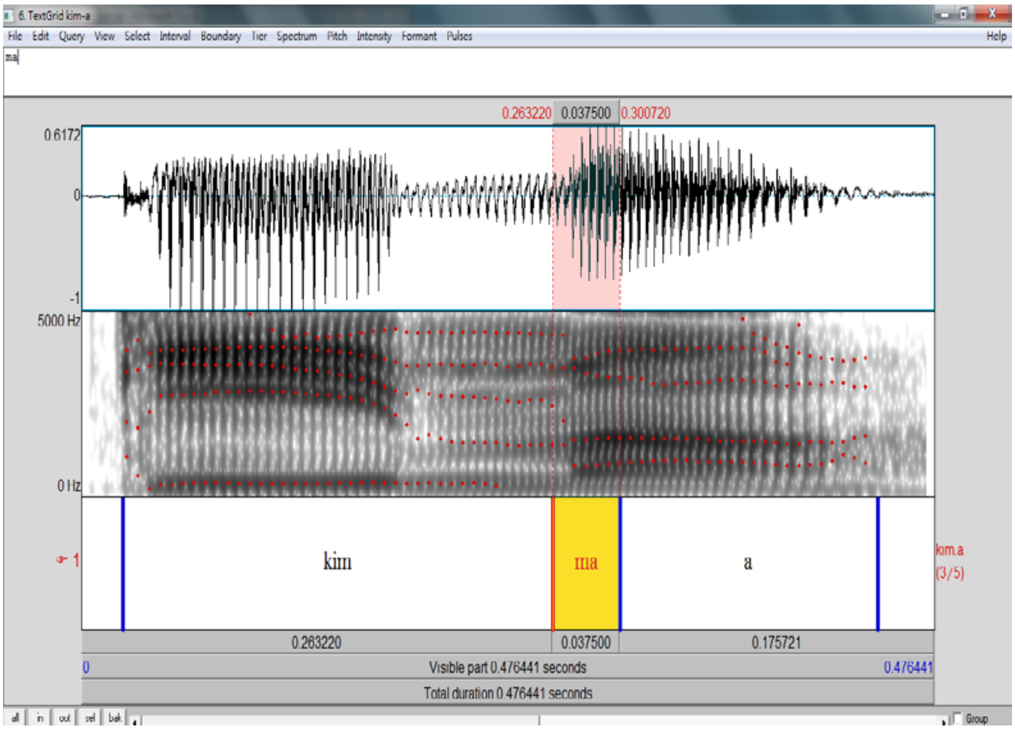
The Word *kim.a* Articulated by Participant #2



The Noticeable Clusters of Frequency Difference Articulated by Participant #2

Time_s	F1_Hz	Frequency difference
0.2058	350.54	N/A
…	…	…
0.2325	386.00	0.04
0.2388	405.69	19.69
0.2450	520.80	115.10
0.2513	600.59	79.79
0.2575	644.27	43.69
0.2638	687.79	43.51
0.2700	695.69	7.90
0.2763	769.34	73.65
0.2825	760.07	-9.26
0.2888	759.87	-0.21
…	…	…
0.4145	490.44	N/A

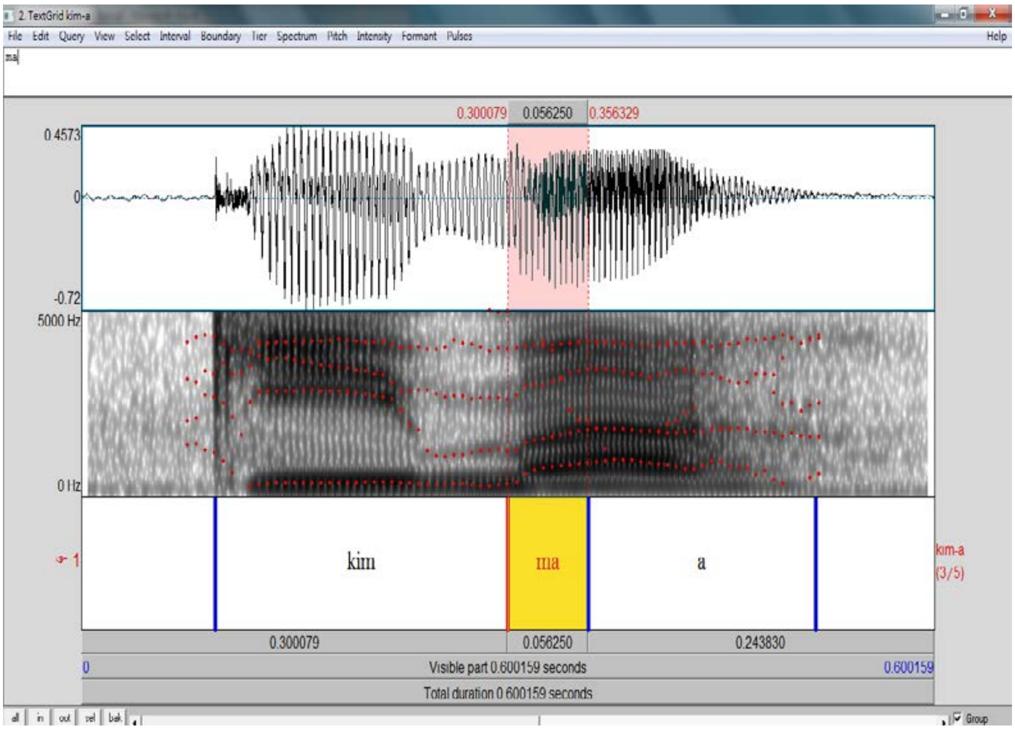
The Word *kim.a* Articulated by Participant #3



The Noticeable Clusters of Frequency Difference Articulated by Participant #3

Time_s	F1_Hz	Frequency difference
0.0231	1038.17	N/A
…	…	…
0.2570	319.91	-1.11
0.2632	337.18	17.28
0.2695	522.72	185.54
0.2757	660.05	137.33
0.2820	689.96	29.91
0.2882	723.77	33.81
0.2945	769.70	45.93
0.3007	806.08	36.38
0.3070	822.19	16.11
0.3132	822.86	0.67
…	…	…
0.4446	828.95	N/A

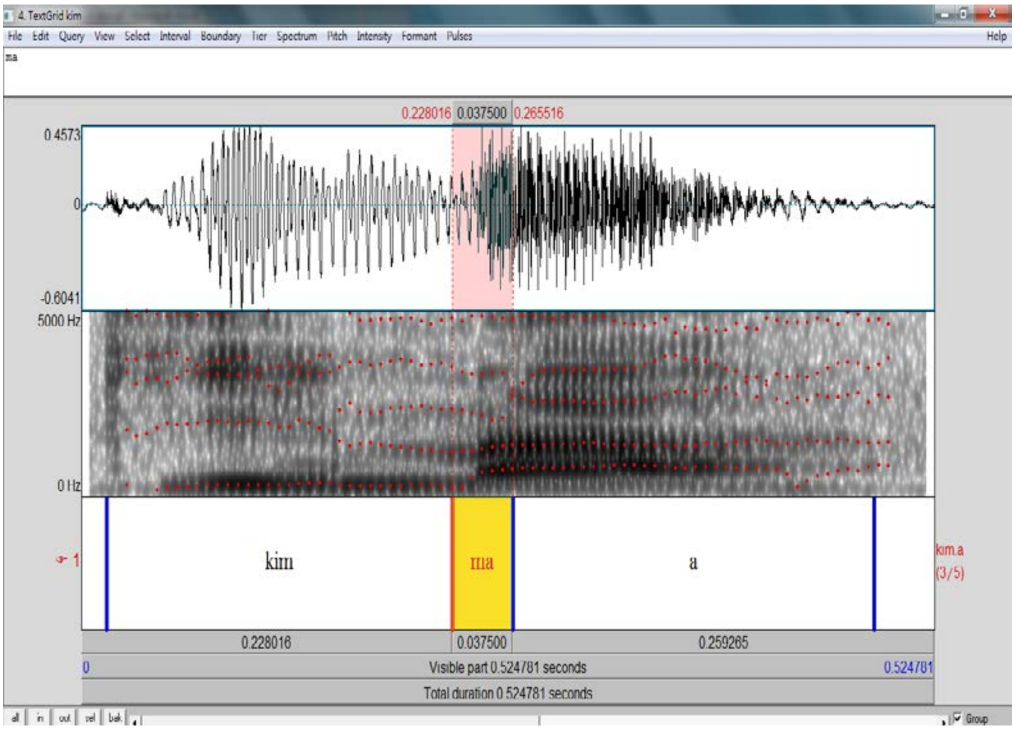
The Word *kim.a* Articulated by Participant #4



The Noticeable Clusters of Frequency Difference Articulated by Participant #4

Time_s	F1_Hz	Frequency difference
0.0941	1145.36	N/A
…	…	…
0.2938	350.30	16.01
0.3001	361.93	11.64
0.3063	386.71	24.77
0.3126	697.14	310.43
0.3188	744.17	47.03
0.3251	764.24	20.07
0.3313	778.89	14.65
0.3376	802.20	23.32
0.3438	841.68	39.48
0.3501	875.32	33.64
0.3563	896.68	21.36
0.3626	856.11	-40.57
0.3688	747.23	-108.88
…	…	…
0.5166	558.66	N/A

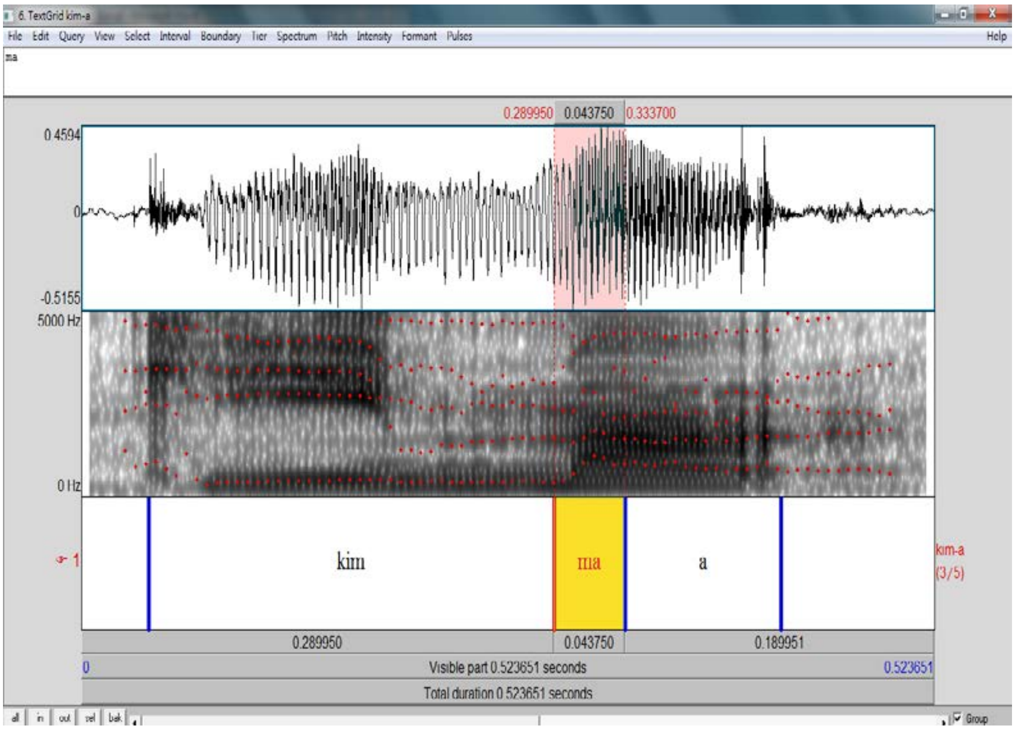
The Word *kim.a* Articulated by Participant #5



The Noticeable Clusters of Frequency Difference Articulated by Participant #5

Time_s	F1_Hz	Frequency difference
0.0157	350.54	N/A
…	…	…
0.2218	263.81	-18.04
0.2280	207.30	-56.51
0.2343	247.76	40.45
0.2405	509.15	261.40
0.2468	637.45	128.30
0.2530	677.53	40.08
0.2593	701.15	23.61
0.2655	729.54	28.39
0.2718	749.53	19.99
0.2780	752.15	2.63
…	…	…
0.4875	808.55	N/A

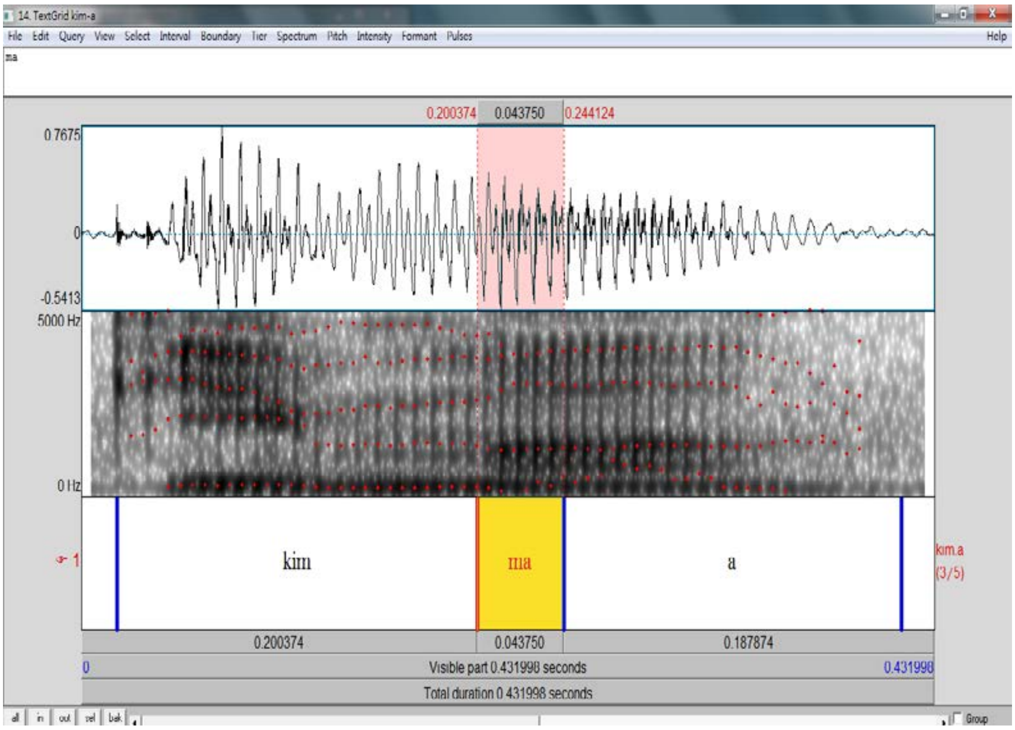
The Word *kim.a* Articulated by Participant #6



The Noticeable Clusters of Frequency Difference Articulated by Participant #6

Time_s	F1_Hz	Frequency difference
0.0412	865.70	N/A
…	…	…
0.2837	458.47	-11.64
0.2900	434.20	-24.27
0.2962	493.16	58.95
0.3025	653.06	159.90
0.3087	782.36	129.30
0.3150	980.84	198.48
0.3212	657.88	-322.96
0.3275	823.45	165.57
0.3337	846.91	23.46
0.3400	835.72	-11.19
0.3462	797.49	-38.23
…	…	…
0.4295	704.48	N/A

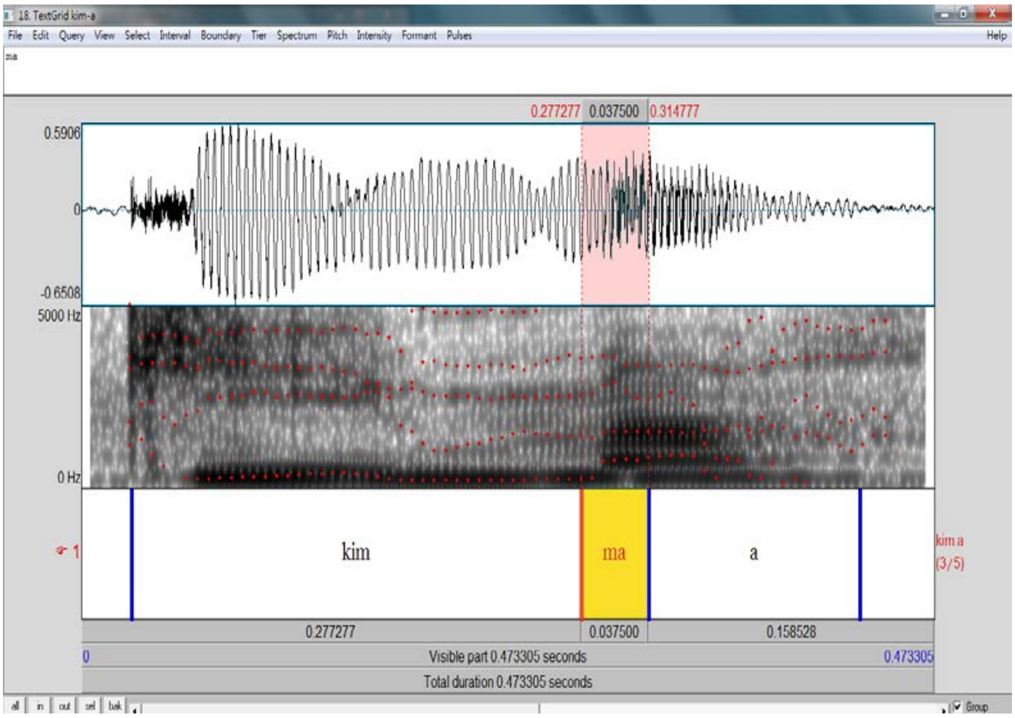
The Word *kim.a* Articulated by Participant #7



The Noticeable Clusters of Frequency Difference Articulated by Participant #7

Time_s	F1_Hz	Frequency difference
0.0180	N/A	N/A
…	…	…
0.1941	226.75	1.54
0.2004	232.36	5.61
0.2066	255.08	22.73
0.2129	294.58	39.50
0.2191	250.30	-44.29
0.2254	826.70	576.40
0.2316	895.47	68.77
0.2379	905.84	10.37
0.2441	932.17	26.34
0.2504	905.29	-26.89
0.2566	903.47	-1.82
…	…	…
0.4150	N/A	N/A

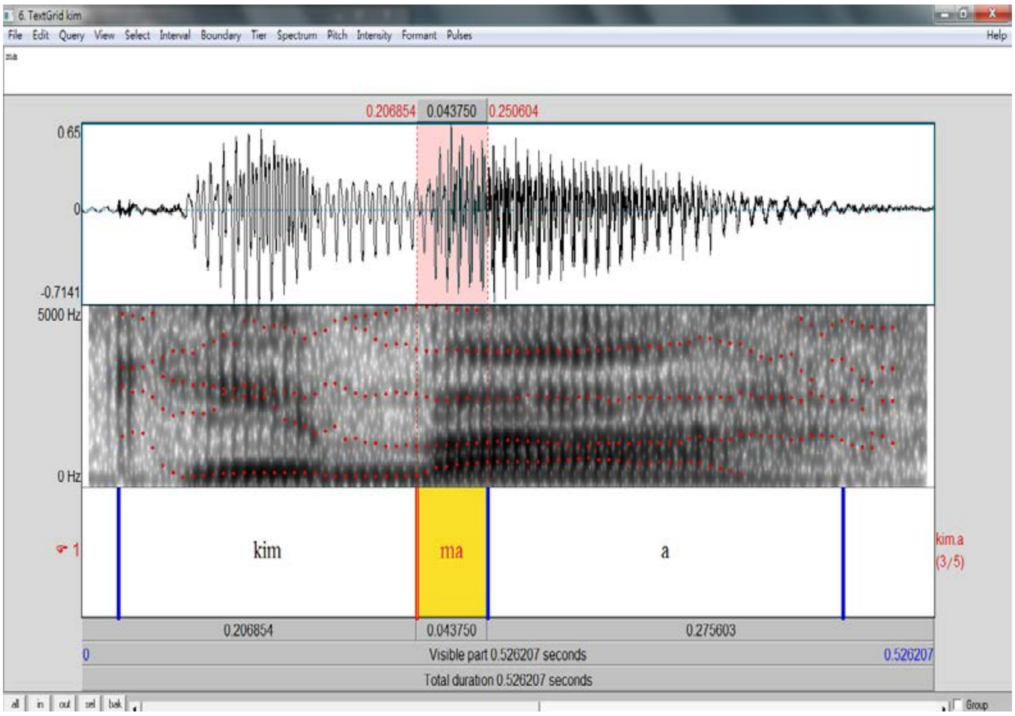
The Word *kim.a* Articulated by Participant #8



The Noticeable Clusters of Frequency Difference Articulated by Participant #8

Time_s	F1_Hz	Frequency difference
0.0277	1167.36	N/A
…	…	…
0.2710	226.67	-0.94
0.2773	232.08	5.41
0.2835	255.88	23.80
0.2898	734.87	478.99
0.2960	772.34	37.47
0.3023	754.54	-17.80
0.3085	784.71	30.17
0.3148	817.91	33.20
0.3210	799.37	-18.54
0.3273	763.45	-35.92
…	…	…
0.4317	1124.71	N/A

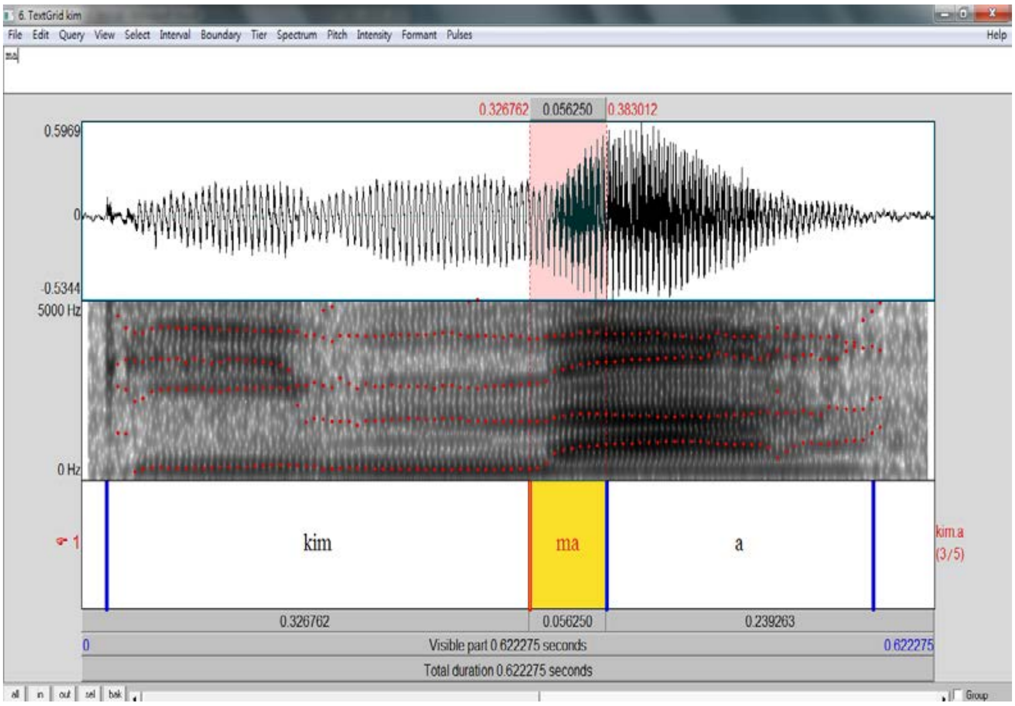
The Word *kim.a* Articulated by Participant #9



The Noticeable Clusters of Frequency Difference Articulated by Participant #9

Time_s	F1_Hz	Frequency difference
0.0230	1379.68	N/A
…	…	…
0.2006	276.30	3.95
0.2069	286.28	9.98
0.2131	406.60	120.32
0.2194	503.21	96.61
0.2256	549.33	46.12
0.2319	579.86	30.53
0.2381	613.34	33.49
0.2444	642.07	28.73
0.2506	663.74	21.67
0.2569	666.65	2.90
0.2631	669.03	2.38
…	…	…
0.4697	1251.87	N/A

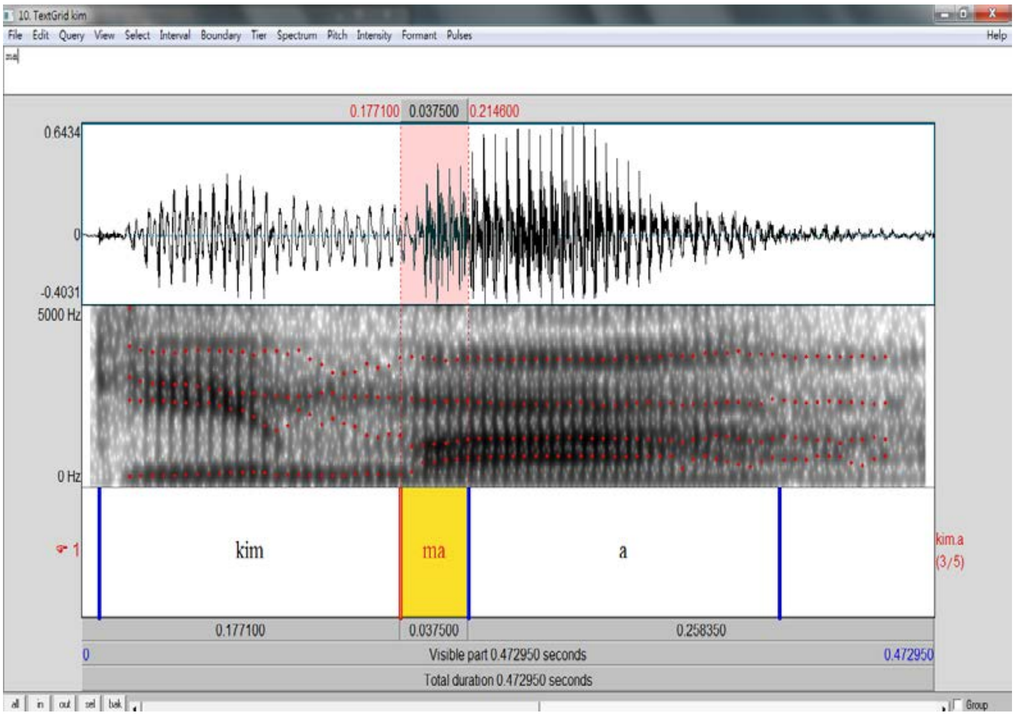
The Word *kim.a* Articulated by Participant #10



The Noticeable Clusters of Frequency Difference Articulated by Participant #10

Time_s	F1_Hz	Frequency difference
0.0186	N/A	N/A
…	…	…
0.3205	286.04	-2.76
0.3268	291.27	5.23
0.3330	312.42	21.15
0.3393	444.69	132.27
0.3455	718.29	273.61
0.3518	802.98	84.69
0.3580	857.39	54.41
0.3643	906.32	48.93
0.3705	927.42	21.10
0.3768	954.04	26.62
0.3830	982.16	28.12
0.3893	989.78	7.62
0.3955	1000.67	10.89
…	…	…
0.5773	1304.77	N/A

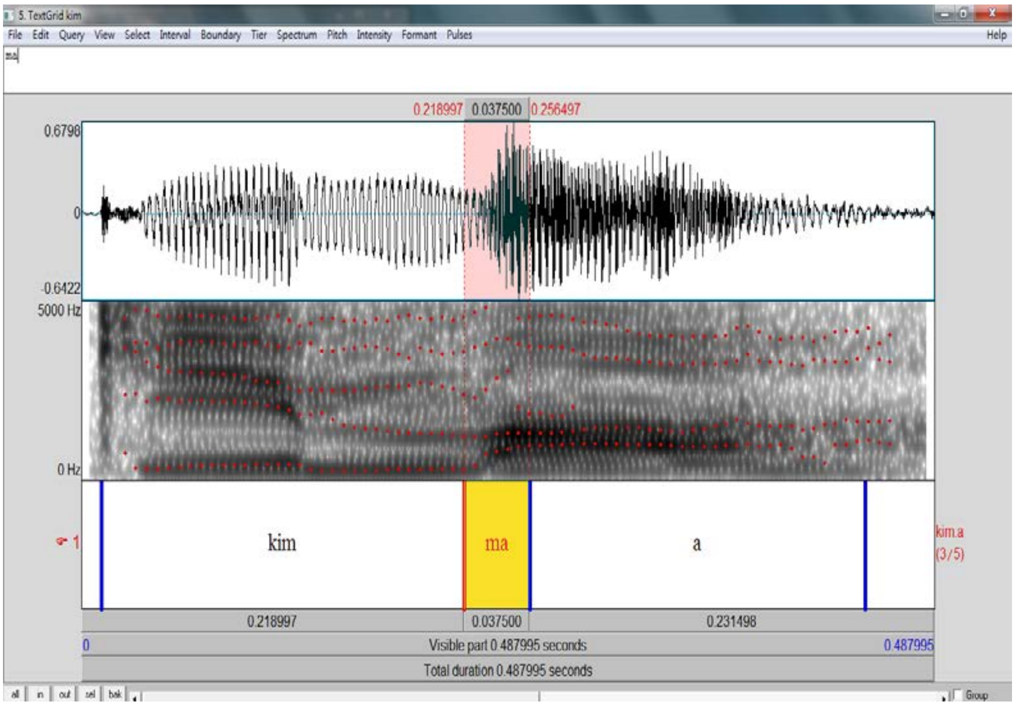
The Word *kim.a* Articulated by Participant #11



The Noticeable Clusters of Frequency Difference Articulated by Participant #11

Time_s	F1_Hz	Frequency difference
0.0098	N/A	N/A
…	…	…
0.1709	309.87	-13.04
0.1771	316.57	6.70
0.1834	390.36	73.79
0.1896	638.09	247.74
0.1959	680.35	42.26
0.2021	715.21	34.86
0.2084	740.22	25.01
0.2146	774.31	34.09
0.2209	788.02	13.71
0.2271	794.08	6.06
…	…	…
0.3867	759.76	N/A

The Word *kim.a* Articulated by Participant #12



The Noticeable Clusters of Frequency Difference Articulated by Participant #12

Time_s	F1_Hz	Frequency difference
0.0113	N/A	N/A
…	…	…
0.2127	283.57	-10.30
0.2190	277.21	-6.36
0.2252	399.64	122.43
0.2315	690.78	291.14
0.2377	784.80	94.02
0.2440	839.38	54.58
0.2502	915.48	76.10
0.2565	953.56	38.08
0.2627	952.50	-1.06
0.2690	942.56	-9.94
…	…	…
0.4481	1080.20	N/A

台語後綴詞「金仔」之過渡音研究

林秉宥

明道大學

本文在自主音段音韻學 (Goldsmith 1976) 的基礎上，提出了「二空缺」的分析架構 (Lin 2006)，並以此分析台語詞彙「金仔」在兩詞素「金」與「仔」結合之後所產生的音韻變化現象 (暫不涉及音節間連音及超音段的分析)。本文提出的過渡音「二空缺」概念，也就是「重複音段」(gemination) 所認為的一個子音空缺，實際上是由二個小音段透過雙向擴散所共同參與的，即 [kim^{ma}a] 而非 [kimma] 。「二空缺」的提出除了有聲學方面的依據外，同時它也較能有效區隔 「金仔」 與 「金馬」 在語音表徵 (phonetic representation) 上的差異，即 [kim^{ma}a] ‘金仔’ vs. [kimma] ‘金馬’，而非 [kimma] ‘金仔’ vs. [kimma] ‘金馬’。

關鍵詞: 詞綴音韻學、自主音段架構、二空缺假設、過渡音、雙向擴散