Phonological Words in Mandarin Speech Production

CHUCK CHENHAO CHIU
National Chung Cheng University

0. Introduction

According to the theory of lexical access in speech production of Levelt, Roelofs, and Meyer (1999), phonological encoding is the stage of constructing metrical frames for phonological words, which, according to them, are crucial units undergoing phonological rules and phonetic executions. Similarly, Levelt (1989, 1992) proposed that the frames of phonological words are basic units in speech production. In his model, the phonological encoding is sensitive to syntactic structure. However, as Nespor and Vogel (1986) and Selkirk (1984) pointed out, prosodic structure and syntactic structure do not have a one-to-one mapping relationship. In this regard, the issue of which constituents, either syntactic or prosodic, are taken as planning units in production becomes interesting. Ferreira (1991) found that sentence length and syntactic complexity affected the preparation time to initiate utterances, but whether phonological factors or the syntactic factors affect the production latency was not distinguished. Wheeldon and Lahiri (1997), following Levelt's model, gave evidence to show that Dutch speakers took phonological words as basic units in speech production. The role of phonological words in prosodic structure for speech production was thus addressed.

Most research on Mandarin prosody has focused on the formation and classification of prosodic constituents or domains and on their relationships with other factors, such as syntactic structures (e.g., Shih 1986, Duanmu 1999, 2000, Lu and Duanmu 2002). These proposals have not addressed the issue of prosodic constituents as planning units in speech production. If Levelt's model is universally true, it should be also applicable to Mandarin Chinese. Following Levelt's (1989, 1992) model and Wheeldon and Lahiri's (1997) findings on Dutch, in the present study I describe two experiments, an experiment of prepared speech and one of immediate (spontaneous) speech to observe the role of phonological words in Mandarin speech production.

1. The Role of Phonological Words

Nespor and Vogel (1986:1) stated that in prosodic theory, the mental representation of speech is divided into hierarchically arranged chunks. As in syntactic structure, lower constituents merge together and become a bigger constituent. In prosodic structure, feet constitute a phonological word, and a number of phonological words form one phonological phrase. They also defined the phonological word as the lowest constituent of the prosodic hierarchy at the level where phonology meets morphology. Nespor and Vogel (1986) proposed that in the typically continuous flow of speech, such mental chunks, the prosodic constituents of the grammar, are signaled by different types of cues ranging from actual segmental modifications to more subtle phonetic changes. That is, the constituents of different domains undergo different phonological rules and phonetic processes. For example, the vowel harmony rule is applied at the domain of the phonological word in Turkish and Hungarian. In English, the phonological phrase mainly has influence on the rhythmic properties and the pause divisions, and the phonological word is the domain of assigning word stress.

Shih (1986) proposed a Foot Formation Rule defining feet and super feet in Mandarin Chinese. However, she did not define whether a foot or a super foot constructs a phonological word. She did not define how many feet there are in a phonological word. Wang (2000) proposed that a Chinese phonological word consists of one foot ((1) and (2)) or two feet (fixed idioms like (3) and (4)).

(1)	*yu3san3*	('umbrella');
(2)	*yu3san3chang3*	('umbrella factory')
(3)	*luan4qi1ba1zao1*	('messy')
(4)	*mei2kai1yian3xiao4*	('joyful appearance')
(5)	*zhi3yu3san3*	('umbrella made of paper')

According to Wang (2000), (5) is regarded as a phonological phrase because the tone 3 sandhi rule first applies to the compound *yu3san3* and later blocks the application of the *zhi3yu3* sequence. Thus, it is considered a phonological phrase, rather than a phonological word. In contrast, the tone 3 sandhi rule applies linearly from left to right in (2), forming one phonological word with the surface *yu2san2chang3*. Moreover, there are also phonological cliticizations found in Mandarin Chinese. Goad, White, and Steele (2003) proposed that a verb inflected by an aspect marker is regarded as one-foot phonological word. As in a phrase like *mai3le0*, the verb *mai3* 'buy' and the aspect marker *le0* form one foot and directly link up to the phonological word. This cliticization is understood as an *internal clitic* as defined by Selkirk (1984).

As for neutral-toned syllables, earlier research provided phonetic evidence and arguments claiming that neutral-toned syllables tend to be cliticized to the neighboring syllables. Chao (1968) pointed out that neutral-toned words tend to have weak stress due to their lower pitch and shorter duration. Hsiao (1991) also claimed that in a presto tempo, function words in Chinese can be assigned leftward to a beat shared with the neighboring content words. These phonetic phenomena support the idea that when an aspect marker happens to be neutral- toned, it is assigned to the preceding verb, forming one phonological domain, specifically the phonological word.

2. The Phonological Word in Speech Production

According to Levelt (1992), in connected speech, speakers do not concatenate all the lexical forms together. The phonological encoding does not derive from the lexical skeleton. The prosody generator, according to the language's specific prosody, puts the phonemic and phonological information into a rhythmic structure across lexical word boundaries. The metrical frames of two lexical forms are first concatenated together to form a new rhythmic frame dependent on the context (i.e., syntactically sensitive). Then the segmental information is associated to the metrical frame by syllabification. These phonological words are the domain of the syllabification and stress assignment. Levelt's model predicts that when we produce sentences, we do not retrieve lexical word forms and concatenate them one by one. We instead construct the metrical frames for phonological words and take these phonological words as basic units in production. Ferreira (1991) found that in structured sentences, the more phonological words included in a sentence, the longer it takes participants to initiate it. Ferreira's findings supported the claim that even though the phonological word is sensitive to the syntactic structure as defined by Levelt (1989, 1992), the effects on production latencies are better explained by prosodic structures, rather than syntactic structures. Similarly, Wheeldon and Lahiri (1997) found that in prepared speech, a sentence containing more phonological words required a longer time for a subject to initiate. In addition, the results from immediate speech showed that production latencies were influenced by the length of the initial phonological word. Dutch was found to have phonological words as basic units in speech production.

Now we turn to the effects on Mandarin speech production. In addition to the foot and phonological word formations in Mandarin (Shih 1986, Wang 2000), Tseng, Huang, and Jeng (1996) claimed that Chinese speakers make use of syllables as segmentation units during speech perception. Nevertheless, no neutral-toned syllables were used as materials in their experiment. The phonological dependency of neutral-toned syllables was not shown by the full-toned syllables. Chen (2000) proposed that the syllable in Mandarin Chinese is represented and processed as a planning unit. In one speech error example, *qing1-du2-du4* (instead of intended *qing1-zhuo2-du4* 'clarity'), the syllable *du* moved to the left and replaced *zhuo*, but the tones remained unchanged. Chen thus claimed that syllables are represented and processed as planning units in Mandarin Chinese. Later, in Chen, Chen and Dell (2002), the implicit priming effects indicated that syllables without tone can function as planning units and that the tones construct the tonal frame for syllables, just as stress constructs the metrical frames in stress languages. These findings provided evidence supporting Levelt's model of the separation between the metrical spellout and the segmental spellout.

However, no similar priming effects on phonological words have previously been tested. Whether the phonological word can be also taken as a planning unit in production has not yet been investigated. The present study consists of two experiments to observe the effect of phonological words on sentence production, and further to reinforce the formation and the application of phonological words proposed by Levelt (1989, 1992).

3. Experiment 1: Prepared Speech Production

3.1. Prediction

The prediction was based on Dutch experiments, e.g. Wheeldon and Lahiri (1997), which predicted that the more phonological words, the longer it takes to initiate utterances.

3.2. Methods

Following the Dutch experiments, the main method was a question-answer procedure. The subjects first saw a phrase and then heard a question referring to the phrase. They were asked to answer the question with the phrase they had seen earlier. The participants were not allowed to respond until they were signaled by the signal beep. The utterance was not simply a repetition of the phrase, since participants had to put the phrase into a complete sentence by adding a subject *ta* 's/he'. Their responses were recorded and measured.

Materials

The materials contained three sentence conditions: Clitic, Nonclitic, and Control. As Goad, White and Steele (2003) pointed out, the aspect marker *le0* is incorporated to the phonological word of the base to which it attaches; clitics that behave in this way are known as an *internal clitics*. In (6), the subject *ta* 'he' is old information provided in the question, carrying no stress or focus. It can also be drawn into a super foot, as Shih (1986) proposed. Consequently, it does not form one independent foot or phonological word. This kind of sentence is categorized as a Clitic sentence, with bound units forming one phonological word.

(6) [*ta* *mai-le*] [*keben*]
he buy-ASP textbook
'He bought a textbook.'

In contrast, the Nonclitic sentence is comprised of words that are not incorporated together into one prosodic constituent, as in (7).

(7) [*ta mai*] [*cuo*] [*keben*]
he buy wrong textbook
'He bought the wrong textbook.'

As Wu (2001) proposed, the resultative word is not cliticized to the verb and thus forms a distinct phonological word (e.g. [*mai* 'buy'][*cuo* 'wrong'] 'bought (something) wrong').

The Control sentences contained the same number of phonological words as the Clitic sentences did but differed in the number of syllables, as shown in (8).

(8) [*ta* *mai*] [*keben*]
he buy textbook
'He buys a textbook.'

Each condition contained four distinct sentences (with different verb-noun combinations), and thus 12 sentence sets were used in total. All sequences causing tone sandhi were avoided.

3.3. Design

All three conditions contained 12 sentences, resulting in 36 target sentences in total. Participants were given Chinese instructions which asked them to respond in a natural speaking rate as soon as they heard the signal beep. They were given demonstration trials and then practice trials. Following the instructions and practice, each participant was asked to produce these 36 target sentences four times each (corresponding to four different preparation times), and thus there were 144 trials in total for each participant. Four preparation times were used to prevent the participants from getting used to the beep pattern. The participants had to wait for three beeps from the offset of the question and then respond. The four kinds of preparation times were produced by manipulating the occurrence of the third beep, set at 750 ms, 1000 ms, 1250 ms, or 1500 ms from the offset of the second beep. The order of target sentences with different preparation times was randomized. Each sentence in all three conditions occurred once at each of the four latencies.

The materials were put into a Latin square design across six blocks. Each target sentence was matched with four preparation times. Two of them occurred in the first three blocks, and the other two occurred in the other half. Each block contained all three kinds of condition sentences, matched with the four preparation times. No sentence occurred more than once within the same block. The number of sentences in each block was equal. The target sentences in each block were randomized to prevent participants from becoming accustomed to the pattern.

3.4. Procedure

All the participants were tested alone in a sound-proof lab. The experiment consisted of instructions, demonstration trials, and practice trials, in that order. The participants were first shown six demonstration trials and then heard recorded example responses. Then they were tested with these same demonstration trials as practice. The trials proceeded as follows. The phrase first appeared at the center of the monitor for 500 ms. After a 500 ms pause, the participant heard a question (which referred to the sentence they saw earlier on the monitor) through the headphones. This question was followed by a series of beep sounds: the first beep sound occurred 2 s from the offset of the question, and the second beep occurred 1 s after the first beep. The third beep was set at one of the four variable time points. Each sentence in all three conditions occurred once at each of the four latencies. There was an approximately 2 second interval between trials. The latencies were measured from the onset of the last beep sound to the voice onset. The response sentences were all recorded. The entire experiment session took around 40 to 50 minutes to complete.

3.5. Subjects

Eighteen student participants (6 males and 12 females) at National Chung Cheng University were tested. They were all native Taiwan Mandarin speakers. None of them was reported to have visual or hearing impairment. They were paid for their participation.

3.6. Results

The analyses were only based on correct response trials. Incorrect responses (e.g., answers were not what they had seen on the monitors) and blank responses were removed from analyses. Responses made before the third beep were also excluded, as well as utterances with any disfluency. Utterances lacking a natural speaking prosodic pattern (e.g. hesitation or particularly slow speech) were also excluded. Responses like the above were all marked as errors. The mean error rate was 4.3%. In addition, responses which exceeded two standard deviations from the mean were also removed. The total rate of data loss was 7.3%. The mean production latencies and error rates across sentence conditions and preparation time types are listed in (9).

(9) Mean latencies (ms) of Experiment 1

Latency (error rate%)				
	Clitic	Nonclitic	Control	Mean
Phonological words	2	3	2	
Lexical words	4	4	3	
Syllables	5	5	4	
Preparation time				
750 ms	472 (6)	483 (5.6)	476 (2.8)	477 (4.8)
1000 ms	462 (4.2)	468 (6)	452 (4.6)	461 (4.9)
1250 ms	454 (3.2)	478 (3.7)	466 (2.8)	466 (3.2)
1500 ms	457 (4.6)	473 (4.2)	464 (3.7)	465 (4.3)
Mean	461 (4.5)	475 (4.9)	465 (3.4)	

A two-way repeated measure ANOVA (sentence conditions * preparation times) was conducted. Both by subject analysis and by item analysis showed significant differences. For the by subject analysis, $F(2,34)=4.917$, MSE= 3952.21, $p=.0133<.05$; for the by item analysis, $F(2,22)=4.49$, MSE=2122.81, $p=$

.0232<.05. Fisher's PLSD showed significant differences between Clitic and Nonclitic sentences (p=.0051<.05), and between Nonclitic sentences and Controls (p=.0027<.05). Tukey/Kramer also showed a significant difference between Clitic and Nonclitic sentences at an alpha level of .05. In addition, as seen in (9), there was no tradeoff between response speed and accuracy. The error rates across the three conditions were not particularly different, whereas response speeds were significantly different.

On the other hand, the ANOVA showed no significant difference across preparation times (by subject $F(3,51)$=1.005, MSE=2637.17, p=.398>.05; by item $F(3,33)$=2.731, MSE=2313.06, p=.0595>.05), and there was no significant interaction of sentence condition and preparation time (by subject $F(6,102)$=.776, MSE=464.36, p=.591>.05; by item $F(6,66)$=.543, MSE=319.64, p=.774>.05).

3.7. Discussion

The results of Experiment 1 were consistent with our expectations. Even though the production latency difference between Clitic and Nonclitic sentences was small (about 13 ms), it was significant. Since Clitic and Nonclitic sentences had the same number of syllables and lexical words, the significant difference between them must have resulted from the number of the phonological words. This indicated that there exist effects of phonological words on Mandarin speech production. The number of phonological words affects the time to produce the sentence. Moreover, as in Dutch, no syllable effect on production latency was found. When given preparation time, only the number of phonological words accounts for the production latency of prepared speech. The number of syllables in a sentence seems not to play an effective role in the stage after the phonological words are constructed.

Given enough preparation time, a speaker constructs the phonological words in a sentence. When these units are completed, the utterance is initiated. However, in daily conversations, our speech productions do not require such long a preparation time. Does the phonological word still affect production without preparation time? Wheeldon and Lahiri (1997) found that once the first phonological word is constructed, the utterance is initiated. The length of the initial phonological word thus affects production latencies. The longer the first phonological word, the more time it takes to produce the sentence. The second experiment in the present study, based on this logic, was carried out to look for the same effect in Mandarin immediate speech production.

4. Experiment 2: Immediate Speech Production

4.1. Prediction

It is predicted that in immediate speech production, sentences with longer initial phonological words should require longer production latencies. With the same materials used in Experiment 1, latencies for Clitic sentences should be longer than those for Nonclitic and Control sentences.

4.2. Methods

The main method was a question-answer procedure, as in Experiment 1. However, no signal beep or additional preparation time was given. The subjects were asked to respond as soon as possible upon hearing the question.

Materials

The materials were those in Experiment 1, except that one sentence pair was removed due to inappropriateness. In addition, as in the Dutch experiment, two filler verbs were designed to match with the noun phrase in each sentence. These filler verbs were all monosyllabic and semantically good in sentences. For example, the two filler verbs to replace *mai* in the sentence *ta mai le keben* 'he bought a textbook' were *jie* 'to borrow' and *dai* 'to bring'.

4.3. Design

All the subjects were given instructions requiring them to answer the question immediately after hearing it. Each of the 72 experimental sentences was produced twice, resulting in a total of 144 sentences in total. The materials were put into a Latin square design across six blocks. Each target sentence occurred once in the first three blocks and once in the last three blocks. The number of sentences in each block was equal. The order of the sentences in each block was randomized.

4.4. Procedure

The procedure was the same as Experiment 1, but without beep latencies. The latencies were measured from the onset of the verb in the question to the voice onset. The response sentences were all recorded. The entire experiment session took around 20-30 minutes to complete.

4.5. Subjects

A new set of 18 native Taiwan Mandarin speakers (6 males and 12 females) was recruited. None of them was reported to have hearing or visual impairment. They were paid for their participation.

4.6. Results

The main data preparation procedure followed that of Experiment 1. The mean error rate was 3.9%. The total rate of data loss was 14.9%. The mean latencies and error rates across sentence conditions are listed in (10).

A one-way repeated measure ANOVA found significant differences (by subject analysis, $F(2,34)=5.4$, MSE=1128.61, $p<.01$; by item analysis, $F(2,46)=4.9$, MSE=1981, $p=.01<.05$). At an alpha level of .05, both by subject Fisher's PLSD and by subject Student-Newman-Keuls showed significant differences between Clitic and Control and between Nonclitic and Control sentences, whereas by subject Tukey/Kramer only found a significant difference between Nonclitic and Control. None of the post-hoc found a significant difference between Clitic and Nonclitic sentences. The error rate of Clitic sentences was particularly raised by two responders who had more errors in Clitic sentences.

(10) Mean Latencies (ms) of Experiment 2

Latency (error rate%)			
	Clitic	Nonclitic	Control
Length (syllables) of initial phonological word	3	2	2
Phonological words	2	3	2
Syllables	5	5	4
Lexical words	3	4	3
Mean latency	840 (5.6)	844 (3.9)	828 (2.3)

4.7. Discussion

If there is an effect of the size of the initial phonological word, the production latencies for Nonclitic sentences and Controls with shorter initial phonological words should be both significantly shorter than those for Clitic sentences. However, contrary to our expectation, the latencies for Nonclitic sentences were not shorter than those for Clitic sentences. ANOVAs with post hoc found a significant effect between Nonclitic sentences and Controls. Crucially, we found that overall responses in our experiments were slower than responses in Dutch (Dutch mean 701 ms; Chinese mean 837 ms). If our responders responded too slowly, they could have used the delay to prepare all the phonological words in the sentence, and the latencies would be affected by the number of phonological words, as we saw in Experiment 1, not by the size of initial phonological word. The overall slow responses thus skewed the pattern from our expectation (i.e., latency for Nonclitic sentences should be significantly different from Clitic sentences). To examine this possibility, the 18 participants were divided into two groups, nine fast responders and nine slow responders. The two groups were divided by their response speeds. If there were an effect of the length of the initial phonological word, the latencies for fast responders should be more Dutch-like (i.e., latencies for Clitic sentences would be significantly longer than those for Nonclitic sentences and Controls, whereas Nonclitic sentences and Controls would have no difference); latencies for slow responders should more be like the pattern in Experiment 1 (in which latencies for Nonclitic sentences were significantly longer than those for Clitic sentences and Controls, whereas the latter two are not distinguished). The results from these two groups were as shown in (11).

(11) Latencies for fast and slower responders

Latency (ms)			
	Clitic	Nonclitic	Control
Fast responders	765	761	749
Slow responders	914	926	907

By subject ANOVA did not find a significant interaction between speed and the three sentence conditions ($F(1,2)=1.6$, MSE=324, $p=.22>.05$). The latencies of fast responders did show a more Dutch-like pattern. The latencies for Clitic sentences were longest, since the initial phonological words in Clitics were longer than those for Nonclitic and Controls. This revealed the effect of length of the first phonological word in immediate speech production. By subject ANOVA showed a nearly significant difference across conditions ($F(2,16)=3.62$, MSE=646, $p=.0505$). Post-hoc analyses found no significance. As for slow responders, their latencies indeed looked more like the pattern in Experiment 1. The latencies for Nonclitic sentences were the longest. By subject analysis also showed a marginal significance ($F(2,16)=3.59$, MSE=807, $p=.0515$). Post-hoc analyses (Fisher's PLSD, Tukey/Kramer, and Student Newman Keuls) found significant differences between Nonclitic sentences and Controls.

The results of slow responders implied that the participants used the delays to prepare all the phonological words in the sentences. Once they had prepared all the phonological words, they started to speak. As for the results of fast responders, Clitic sentences containing longer initial phonological words required more time to initiate the utterances. An effect of the size of an initial phonological word on immediate production latencies was thus found. Nonclitic and Controls did not show any difference, as we expected. No interaction between speed and sentence conditions implied that there may be more than one effect affecting latencies. When initiating an utterance, the length of the first phonological word and the total number of phonological words may both affect the production latency. From the current experiments, we don't know whether these two factors can be completely distinguished during sentence production. One thing we can be sure of is that the longer the participants prepared for utterances, the more crucial the total number of phonological words was. The results of Experiment 2 let us conclude that the production latency in immediate speech can be affected by the size of the first phonological word, the total number of phonological words, and the responding speed.

5. General Discussion

Levelt (1989, 1992) proposed a model of phonological encoding claiming that speech production is a process of putting lexical concepts into phonetic articulation. During the phonological encoding, a speaker constructs the metrical frames for phonological words by constructing the foot structures and stress patterns of neighboring lexical items. These phonological words are basic units to produce sentences, as shown in Dutch.

The finding in Experiment 1 implied that Mandarin speakers do construct metrical frames for phonological words and that these phonological words are the basic units in speech production. Thus, the number of phonological words affects the production latency. Levelt's (1989, 1992) model predicted that to initiate an utterance, at least the first phonological word should be constructed. The overall slow responses in Experiment 2 made the result similar to the pattern in Experiment 1. The speakers prepared all the phonological words in the sentences during their longer latencies. The effect of length of the initial phonological word was thus hidden. Nevertheless, we did find a significant effect of size of the phonological word from the fast responders. If the Mandarin speakers generally produce sentences slower than the Dutch speakers, it could be understood as showing that the mixed effects of phonological word number and the size of first phonological word both influence the production latency. In brief, our results corresponded to Levelt's (1989, 1992) model and Wheeldon and Lahiri's (1997) findings that the phonological words are the basic units in speech production. The production latencies could not be accounted for by the number of syllables even though syllables were the phonetic execution units in production. Additionally, the overall slow responses, especially for Nonclitic sentences, in immediate production could be also due to the mixed effects of the number of phonological words and of the size of the initial phonological word. Not only the number of phonological words but also the size of initial phonological word affects the production latencies. With preparation time, the number of phonological words determined the production latency. In immediate speech production, the number of phonological words and the length of first phonological word were influential to the production latencies. The more phonological words or the longer the initial phonological word, the longer the latencies.

Furthermore, it was shown that the production latency is not solely determined by one or two factors. The total number of phonological words, the sizes of the first phonological words, and the speaking rate have interaction effects on production. They all influence the production latencies. At different response rates, the effective factors would be different. In addition, different effects are found at different times in production. Even though in the present study, we did not conclude which factor has effects on the production latency at which stage, or whether these effects occur simultaneously or sequentially, one can bear in mind that the prosodic structure indeed affects Mandarin speech production.

References

Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.

Chen, J-Y., T-M. Chen, and G.. S. Dell. 2002. Word-form encoding in Mandarin Chinese as assessed by the implicit priming task. *Journal of Memory and Language* 46:751-781.

Chen, Jenn-Yeu. 2000. Syllable errors from naturalistic slips of the tongue in Mandarin Chinese. *Psychologia* 43:15-26.

Duanmu, San. 1999. Metrical structure and tone: Evidence from Mandarin and Shanghai. *Journal of East Asian Linguistics* 8:1-38.

--. 2000. Hanyu de jiezou (Prosody in Chinese). *Dangdai Yuyanxue* (*Contemporary Linguistics*) 4:203-209.

Ferreira, Fernanda. 1991. Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language* 30:210-233.
Goad, H., L. White, and J. Steele. 2003. Missing inflection in L2 acquisition: Defective syntax or L1 constrained prosodic representations? *Journal of Canadian Linguistics* 48:243-263.
Hsiao, Yuchau. 1991. *Syntax, rhythm and tone: A triangular relationship*. Ms., University of California, San Diego.
Levelt, Willem J. M. 1989. *Speaking: From Intentions to Articulation*. Cambridge, MA: MIT Press.
--. 1992. Accessing words in speech production: Stages, processes and representations. *Cognition* 42:1-22.
Levelt, Willem. J. M., Ardi Roelofs, and Antje S. Meyer. 1999. A theory of lexical access in speech production. *Behavioral and Brain Science* 22:1-75.
Lu, Bingfu, and San Duanmu. 2002. Rhythm and syntax in Chinese: A case study. *Journal of the Chinese Language Teachers Association*. 37(2):123-136.
Nespor, Marina, and Irene Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris.
Selkirk, Elisabeth O. 1984. *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge: The MIT Press.
Shih, Chi-Lin. 1986. *The Prosodic Domain of Tone Sandhi in Chinese*. Ph.D diss., University of California, San Diego. Ann Arbor: University Microfilms International.
Tseng, C-H., K-Y Huang, and J-Y Jeng. 1996. The role of the syllable in perceiving spoken Chinese. *Proceedings of the National Science Council* 6(1):71-86.
Wang, Hongjun. 2000. Chinese prosodic word and prosodic phrase. *Zhongguo Yuwen* 6:525-536.
Wheeldon, Linda, and Aditi Lahiri. 1997. Prosodic units in speech production. *Journal of Memory and Language* 37:356-381.
Wu, Ching-Huei. 2002. On *de/bu* and the syntactic nature of resultative verbal compounding. Ms., McGrill University.

Chuck Chenhao Chiu
20, Ln 106, Hsi Kun 2nd St.
PanChiao City
Taipei County
Taiwan

chjacques@yahoo.com