篇名：
Productivity of Different Tonal Overwriting in Taiwan Mandarin Cute Talk Reduplication

作者:
陳立信。康橋國際學校。高二C班

指導老師:
黃子權 老師
周德貞 老師

I. Introduction

1. Background and Motivation

As a tone language, Taiwan Mandarin (TM) has 4 basic tones–high [H] (T1), rising [LH] (T2), dipping [L] (T3), and falling [HL] (T4). These tones are morphological, meaning that they convey different meanings when added to Chinese characters. Further, there is tonal processing in TM, including tonal overwriting and T3 sandi (Zhang, 2007). The morpho-phonological tonal overwriting occurs in TM: one common situation is in "cute-talks," which comes with nominal reduplication often, dubed cute talk reduplication (CTR) (Huang, 2016).

CTR is not a popular topic in the linguistics field, hence it is seldomly being researched on. The only person that did something related was T. C. Huang from National ChengChi University. He investigated the productivity of different tone patterns of tonal overwriting in CTR through conducting a survey-based experiment. His experiment, however, could be improved. I discovered the design of his wug words was not rigorous enough: they are subjective and introspectively determined. Further, his choice of the bases were so common that lexical factors were not filtered completely, which lexical variations and factors could not really be exempt. Based on the experimental gaps I spotted in Huang's experiment, I decided to conduct an experiment to reexamine the productivity of different tonal patterns in CTR.

2. Research Question

Lexical factors are shown to be highly associated with the outcome of tonal patterns (Huang, 2016). My research questions are focused on the relationship between lexical factors and the tonal overwriting patterns.

    a. Are the cute-talk fixed tonal patterns psychologically real?
    b. Which pattern is more productive?
    c. How do lexical factors shape and determine the choices of the patterns?

3. Research Purpose

CTR is not a common topic in the linguistic field and previous studies were not necessarily complete and could be improved nonetheless. This paper is to examine three main things: the psychological reality of cute talk fixed tonal patterns [T3-T1] and [T3-T2], the productivity of different tonal patterns in CTR, and the relationship between the degree of lexical factors and the choice of the tonal patterns from native TM speakers.

II. Literature Review

1. HSU, HUI-CHUAN, Zailun Hanyu de Shengdiao yu Zhongyin

Hsu (2006) had remarks on CTR tone patterns of kinship terms. She came out with several patterns and restrictions on kinship term reduplications. First of all, she states that all non [L] bases would end up with reduplications with [L-LH] tone patterns. There is a restriction with all [L] bases, however. She believes that the [L-LH] pattern fails to apply to [L] bases. She found a fixed tonal pattern [L-H], instead as an alternative, for all [L] tone bases.

(1) Examples of fixed tonal patterns (adapted from Hsu, 2006)

a. [L-LH]: ma^H 'mother' → ma^L-ma^LH
ba^H 'father' → ba^L-ba^LH
shu^LH 'uncle' → shu^L-shu^LH
b. [L-H]: jie^L 'sister' → jie^L-jie^H (vs. *jie^L-jie^LH)
gou^L 'dog' → gou^L-gou^H (vs. *gou^L-gou^LH)

As can be seen in (1), jie and gou with [L] tones both have [L-H] tones since they fail to work with [L-LH] tone patterns.

Hsu (2006) provides 2 tonal patterns that appear in TM CTR: [L-LH] and [L-H]. My assumption, however, clashes with Hsu's remarks. I believe that both [L-LH] and [L-H] work with [L] tone bases. Moreover, I believe that the tonal overwriting does not only apply to kinship terms; I believe this tonal overwriting could be applied to all nominal reduplications, including objects.

2. YEH CHIA-HSIN, Comparison of Phonetic Naturalness between Rising-Falling and Falling-Rising Tonal Patterns in Taiwan Mandarin.

Yeh (2010) believes that these CTR tonal overwriting are not restricted to kinship terms, but are restricted by the original base tones. He noted that while [L] and [LH] bases can have reduplications with [L-LH], [H] and [HL] tones fail to work with the [L-LH] tonal overwriting. Table (2) demonstrates this restriction through examples.

(2) Examples and Restriction of the [L-LH] pattern (adapted from Yeh, 2010)

Bases	Gloss	Normal	The pattern applied
$ding^\text{H}$	'name'	$ding^\text{H}$-$ding$	*$ding^\text{L}$-$ding^\text{LH}$
wa^LH	'doll'	wa^LH-wa	wa^L-wa^LH
ni^LH	'name'	ni^LH-ni	ni^L-ni^LH
bao^L	'baby'	bao^L-bao	bao^L-bao^LH
wei^L	'name'	wei^L-wei	wei^L-wei^LH
li^HL	'name'	li^HL-li	*li^L-li^LH
dan^HL	'egg'	dan^HL-dan	?dan^L-dan^LH

Yeh provides base tone restrictions in CTR. While I agree with Yeh's remarks on tonal overwriting and restrictions based on base tones, I disagree with his tonal overwriting patterns. While he noted that the [L-LH] tone can only be used, I believe that both [L-LH] and [L-H] tonal patterns should be allowed since both are common to native TM speakers.

3. HUANG TZU-CHUAN , I-R Tonal Mapping: Evidence from the "Cute-talk" Reduplication in Taiwan Mandarin

Huang (2016) argued and proved that tonal overwriting is morpho-phonological instead of phonological. First, the tonal shift has a sociopragmatic function–the different tonal shift could change a word's meaning. For example, as Huang provided, with the same stem shui, meaning 'water,' 'shui$^\text{LH(T2)}$-shui$^\text{L(T3)}$' and 'shui$^\text{L(T3)}$- shui$^\text{LH(T1)}$' has different meanings; one means 'beauty' and the other means 'water in puerile style.' Second, the tonal patterns can be applied to non-reduplications in a few cases, such as 'a$^\text{L(T3)}$-yi$^\text{LH(T2)}$' 'aunt' and 'bei$^\text{L(T3)}$-bi$^\text{LH(T2)}$' 'darling,' both becoming "cute talks" and convey a connotation of a close family member relationship. This tonal shift has a similar effect to adding an affix. The proven morpho-phonological identity of tonal overwriting allows me to make sure that the tonal overwritings could be applied on other nominal reduplications besides kinship terms.

Huang (2016) also conducted an experiment that triggered me to reexamine his results. He tested the productivity and psychological reality of different CTR tonal patterns with common roots and wug tests. His results turned out that [T2-T3] is the most common and [T3-T1] is more common compared to [T3-T2] as the lexical factor decreases. However, the design of his wug words was not rigorous enough, which led to experimental errors and his goals, lexical variations, and factors could not really be exempted by his experimental design.

4. ZHANG JIE & LAI YU-WEN, Two aspects of productivity in Taiwanese Double Reduplication

Zhang and Lai (2007) also conducted an experiment on reduplication but in Taiwanese. Their experiments were more carefully-done, which I followed their procedure and experimental designs when I conducted my own experiment. In their experiment Zhang and Lai also applied nonce words and wug tests and designed 2 types of wug words: non-reduplicating bases and bases that are accidental gaps. Their base categories are much finer-grained and have a clearer effect on results by lexical factors.

III. Research Methods

1. Participants

My classmate Jeremy and I conducted an experiment on TM's CTR. Our participants are randomly selected from 10th grade of Kang Chiao High School in 2021. There are 4 males and 4 females in total, each taking up 50% of the population. Although some stayed abroad, all of our participants are TM native speakers and have Taiwanese parents. All of our participants stayed in Taiwan for at least 5 years. Besides being native speakers, some of the participants are bilingual and some trilingual: languages/dialects include TM, Japanese, English, Taiwanese, French, Spanish, Cantonese, Henan-Mandarin, and Hakka.

2. Materials

In order to reduce the experimental error (e.g. participants' associations with everyday speech) and see how the degree of nonce word would affect the participants, we designed 4 types of reduplication and a total of 19 of them as our test items. All of them were formed with T3 bases. Below are the four types.

(a) Type 1 (equivalent to Zhang's Type 2) reduplication are formed with bases that are common, high-frequency real characters and have bounded roots, meaning that the characters are part of words and could not stand alone (Zhang, 2007). These are the control group and have the greatest lexical factor.

Example 1	Example 2	Example 3	Example 4
"蚤"	”藻“	"影“	“嶼”

(b) Type 2 are bases formed with uncommon, low frequency real monosyllabic characters. Though they might be bound or free roots, they rarely occur and are non-vernacular to native speakers. The characters are also dead and archaic. Before the experiment, we made sure none of the participants had seen the characters before, and thus they required the Chinese Phonetic Alphabet to read the words during the experiment. These characters have the second greatest lexical factor.

Example 1	Example 2	Example 3	Example 4	Example 5
"鮡(ㄓㄠˋ)"	”卣(ㄧㄡˇ)“	”魃(ㄎㄨㄟˋ)”	“蒹(ㄏㄢˇ)”	“酐(ㄏㄤˇ)”

(c) Type 3 (equivalent to Zhang's Type 3) are monosyllabic bases formed with Chinese Phonetic Alphabets only (Zhang, 2007). The given Chinese Phonetic Alphabets fail to correspond to real nominal character, for example, ㄩㄥˇ in table below; it can only correspond to some existing verbal root, such as 湧, or adj root, such as 勇. In other words, in cute talk reduplication, which is usually formed with nominal bases, type 3 are quasi-accidental gaps, which, for a syllable to count as an accidental gap, the segmentals and the tone must both be valid for TM, but their combination happens to be missing from the TM syllabary (Zhang, 2007). These bases have minimal lexical factors. They are not shown in any forms of Mandarin characters so the participants, when encountering these bases, would have the least association with other words/characters, reducing the lexical factors of these bases.

Example 1	Example 2	Example 3	Example 4	Example 5
"ㄈㄡˇ"	“ㄌㄨㄥˇ”	"ㄌㄡˇ"	"ㄆㄡˇ"	"ㄩㄥˇ”

(d) Type 4 are bases phonotactically ungrammatical in Mandarin, meaning that they are systematic gaps and could not be written in TM characters nor in Chinese Phonetic Alphabets. Therefore these bases are given in Romanization. These bases were inspired by Hakka when designing so some of them are transliterations. Type 4 has no lexical factor. They are not shown in any forms of Mandarin characters so the participants, when encountering these bases, would have no association with other words/characters, reducing the lexical factors of these bases.

Example 1	Example 2	Example 3	Example 4	Example 5
"Hiet3"	"Yamn3"	"Sam3“	"Ngu3"	"Heu3"

These 4 types of mono-syllabic bases are placed between quotation marks like "藻" in (3). They are not shown in reduplicated form because we do not want the participants to be distracted by the already formed reduplication. We do also want the participants to go through the processing step. These 19 bases are then used to create 19 sentences in order to provide the participants scenarios of cute-talking. These 19 sentences are then randomized to distract the participants, preventing them from noticing the experimental content, reducing experimental errors.

(3)
　　　　　　海裡的 ”藻“ 好綠, 好像油畫裡會出現的喔

We also have fillers in our materials. Similar to Huang's (2016) fillier in his experiment, we placed several character bases with different tones (T1,T2,T4) in our sentence list. This avoids the participants from realizing that characters all have the same tone T3.

Another piece of material is designed to be our training trial for our participants. The structure of the training trial is the same (CF. Huang, 2016). The difference between the training trial and the formal trial is the bases and the sentences they are put into. The bases are non T3 tones and the sentences are similar but not identical to the ones in the formal trial.

3. Procedure

Before both the training and formal trials, we inform the participants to imagine talking to a lover/baby/pet in a cute way. Every subject then took both the training trial and the formal trial. The training trial was to ensure the participants could perform the experiment correctly before the formal trial. During the training trial, the participants were instructed to reduplicate the monosyllabic bases in the quotation marks to form a reduplicated word, and read out all the sentences from the sentence list we provided. We made sure our participants could read out the sentences as we expected before they took their formal trial. When there was a problem in the participant's training trial, they would be excluded from the formal trial.

During the formal trial, we provided our participants a sentence list with 4 types of TM T3 reduplications (see materials section) and a total of 19 target sentences for each participant. Participants read our 19 target sentences and filter sentences directly and recorded their voices (readings) for our analysis.

During the trials, we informed our participants that type 3 and 4 words were taken from other languages/dialects to prevent and reduce effects from lexical factors and to prevent awkwardness in performing the trial in order to have the participants remain in cute-talking. This ensured that our experiments were consistent.

4. hypothesis

I hypothesize that the more real a base is, the greater the lexical factor is, the more frequent the [T3-T2] and [T3-T1] tone pattern would appear. Because [T3-T2] and [T3-T1] are commonly used in native TM everyday conversations (Huang, 2016), the greater the lexical factor of the base is, the greater association (Bybee, 2002) to the mental lexicon the participants would have due to the priming effect (Cherry, 2021). In other words, the faker the base is, the lower the lexical factor is, CTR tonal patterns would not be as productive. Therefore no CTR tonal patterns would be used and the norm [T2-T3] would become more frequent. [T2-T3] is considered as the norm because T3 sandhi ([T2-T3]) is strongly planted in native TM speakers' cognition (Yeh, 2010).

Between [T3-T2] and [T3-T1] CTR tonal patterns, I hypothesize [T3-T1] would be the most effective out of the two. This hypothesis is formed based on the results of Huang's (2016) experiment, which showed higher frequency in the use of [T3-T1] compared to the use of [T3-T2] of his participants.

Based on the above, in short, I hypothesize [T3-T2] and [T3-T1] occur most frequently in Type 1 bases and least frequent in type 4 bases. Between [T3-T2] and [T3-T1] CTR tonal patterns, [T3-T1] would be the most effective out of the two. [T2-T3] occur most frequently when the lexical factor of a base is the least.
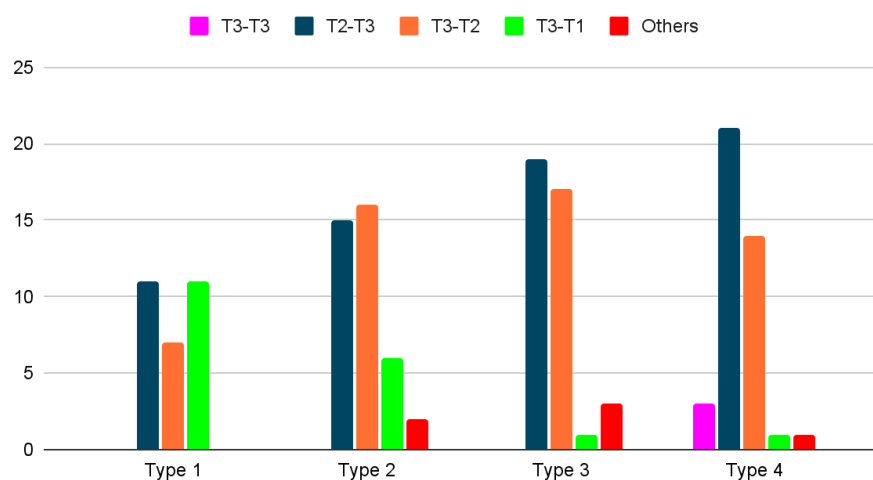
IV. Analysis and Results
1. Raw Data Table

Table 1:

	Type 1	Type 2	Type 3	Type 4	Total
[T3-T3]	0 (0.00%)	0 (0.00%)	0 (0.00%)	3 (7.50%)	3 (2.03%)
[T2-T3]	11 (37.93%)	15 (38.46%)	19 (47.50%)	21 (52.50%)	66 (44.59%)

[T3-T2]	7 (24.14%)	16 (41.03%)	17 (42.50%)	14 (35.00%)	54 (36.49%)
[T3-T1]	11 (37.93%)	6 (15.38%)	1 (2.50%)	1 (2.50%)	19 (12.84%)
Others	0 (0.00%)	2 (5.12%)	3 (7.50%)	1 (2.50%)	6 (4.05%)
Total	29	39	40	40	148

Graph 1:

Bar Graph of Types of Bases vs Tonal Patterns

2. Results

   a. [T3-T3]
      In CTR, [T3-T3] tone pattern takes up a small percentage out of all tone patterns, only a total of 2.03%. [T3-T3] also only occurs in type 4 bases, which is a base type with a small lexical factor. There is an upward trend in the use of it.

   b. [T2-T3]
      In CTR, [T2-T3] tone pattern takes up the greatest percentage out of all tonal patterns, with a total of 44.59%. [T2-T3] tone pattern is also one of the highest in each individual type of bases. [T2-T3] shows a clear linear upward trend as the lexical factor decreases.

   c. [T3-T2]
      In CTR, [T3-T2] tone pattern takes up a great percentage of all tone patterns, with a total of 36.49%, coming after [T2-T3]. It is greater than the [T3-T1] tone pattern. Like Other tone patterns (red), [T3-T2] has its highest at type 2 and 3, sharing the same trend as other tone patterns (red). [T3-T2] tone pattern shows a more average data, with just a slight rising slope in the trend as lexical factor decreases.

   d. [T3-T1]
      In CTR, [T3-T1] tone pattern takes up only 12.84% of all tone patterns. [T3-T1] shows a significant drop in its frequency as lexical factor decreases. Initially, in type 1, it was the most frequently used tone pattern. It then drops to the least frequently used at the end, in type 3 and 4.
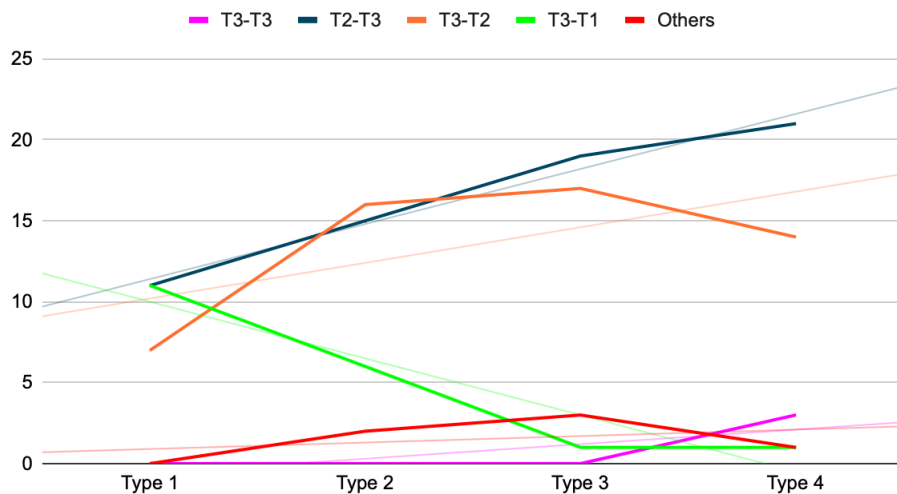
e. Other tone patterns

In CTR, other tone patterns (red) take up the least percentage of all tone patterns, with only 4.05%. Another tone pattern shares a similar trend as [T3-T2] tone pattern has. They peak at type 2 and 3, show more average data, and have a slight rising slope in the trend as lexical factor decreases. Other tone patterns do not appear in type 1 bases, but do appear in types 2-4.

As lexical factor decreases, the difference between the frequency of the different tonal patterns diverges. [T2-T3] shows an upward trend, [T3-T1] shows a downward trend, and the rest has a comparatively insignificant change in the data. While having comparatively insignificant change in the data, [T3-T2] and others tone pattern (red) still peaks in type 2 and 3 and is at the lowest point in type 1 and 4.

3. Analysis

Graph 2:

Line graph of Types of Bases vs Tonal Patterns

a. [T3-T3]

[T3-T3] tone pattern takes up a small percentage out of all tone patterns. This is expected as this tone pattern is not common in TM due to the existence of T3 sandhi (Yeh, 2010). As a T3 follows a T3 base, T3 sandhi, a common rule strongly planted in native TM speakers' cognition, would be activated as the participant recognizes the language as TM (Yeh, 2010). In this experiment, our bases, including wug words, are all designed and chosen from TM and close dialects. This explains why this tone pattern takes up a small percentage out of all tone patterns since most [T3-T3] were transformed. The only instance [T3-T3] occurs is in type 4 bases, which is a base type with systematic gaps. This is also understandable because this is the type that differs the most from TM. Some participants would not have recognized this type of base as TM in their cognition and thus T3 sandhi does not take place (Yeh, 2010).

b. [T2-T3]

Matching my hypothesis, there is an increasing trend in the [T2-T3] tone pattern as the type of bases decrease in their lexical factors. As shown in graph 1, the line of best fit has a steady positive linear slope, indicating that there is a directly proportional relationship between the use of [T2-T3] tone pattern and the degree of lexical factors of the bases: the greater the lexical factor of a base is, the less the frequent [T2-T3] tone

pattern, the normal tone pattern, is used. Since [T2-T3] is considered as the norm due to the fact that T3 sandhi ([T2-T3]) is strongly planted in native TM speakers' cognition (Yeh, 2010), the original [T3-T3] would be replaced by [T2-T3] everytime a double T3 shows up. The participants used [T2-T3], the "normal tone," more in type 3 and 4 bases because in their cognition, the cute-talk tonal patterns [T3-T1] and [T3-T2] were not that productive and did not fit the scenario, regardless to the fact that tone pattern [T3-T2] and [T3-T1] exists in their cognition. The connection to their mental lexicon weakens and the priming effect (Cherry, 2021) was not as effective as the words became unreal and the lexical factors decreased. This means that using the CTR tone patterns [T3-T2] and [T3-T1] is not as productive and has no psychological reality in that context.

   c. [T3-T2]

      [T3-T2] tone pattern takes up a great percentage of all tone patterns (25%), becoming more frequently used compared to [T3-T1] CTR tone pattern. This tonal pattern is more frequently used because it appears more common in everyday conversation since, as Hsu (2006) points out, the [T3-T2] tone pattern can also apply to non-T3 bases. Therefore, the participants would have more exposure in [T3-T2] tone patterns in their everyday lives and would have more connections to their mental lexicon and have stronger priming effects (Cherry, 2021) when seeing these types of bases (Bybee, 2002). Furthermore, [T3-T2] is at its peak at type 2 and 3 because type 2 and 3 are bases with the least lexical factors while they are bases that clearly convey a sense of TM to the participants since they are both shown in or with Chinese Phonetic Alphabets. In addition, [T3-T2] tone pattern shows a more average data, with just a slight rising slope in the trend as lexical factor decreases. This indicates that the [T3-T2] tone pattern is effective enough for no great loss in the population who chose to use it even though there is a change in lexical factors.

   d. [T3-T1]

      The results of the [T3-T1] tone pattern clashes with both parts of my hypothesis and Huang's (2016) result that [T3-T1] tone pattern would be more effective compared to [T3-T2] tone pattern. [T3-T1] tone pattern takes up only 12.84% of all tone patterns, coming below [T3-T2] tone pattern. Moreover, [T3-T1] tone pattern shows a significant drop in its frequency as lexical factor decreases, which is as expected in the hypothesis. These outcomes above all prove that the [T3-T1] tone pattern is ineffective for T3 CTR since the frequency did not maintain when the lexical factor decreases.

   e. Other tone patterns

      The other tone patterns appear in types 2, 3, and 4 only. This effectively support that the participants have strong associations to their mental lexicon and that priming effects (Cherry, 2021) were effective in type 1 bases, thus only the normal tone pattern [T2-T3] and CTR tone patterns [T3-T1] and [T3-T2] are used in the category with the least lexical factor. The use of other tone patterns in type 2, 3, and 4 is because types 2, 3, and 4 are types that do not occur in regular CTR, and thus a part of our population would feel uncomfortable using CTR tone patterns and would use other tone patterns instead. These tone patterns might be associated to their own experience of language use (ie. they might have their own special associations with other words when they see our designed bases).

V. Conclusion and Suggestions

   1. Conclusion

As seen in the data, [T2-T3] takes up a greater percentage of all tonal patterns and the frequency of the use of [T2-T3] increases significantly while the lexical factor of the bases decreases. Participants chose to use non-CTR tonal patterns [T2-T3] more frequently instead of CTR tonal patterns [T3-T1] and [T3-T2] in all 4 types. Therefore CTR tonal patterns [T3-T1] and [T3-T2] are not psychologically real nor productive. It is shown that the [T2-T3] tone pattern is the most productive out of all other patterns, and between the CTR tonal patterns, [T3-T2] is much more productive than [T3-T1] since [T3-T2] takes up a greater percentage and the frequency difference between the two tonal patterns diverges at the end, which [T3-T2] was much more frequently used. My results differ from Huang's (2016) experiment result, as Huang concludes that [T3-T1] is much more effective than [T3-T2]. The Lexical factors had a huge impact on the choice of tonal patterns. As results have shown, the greater the lexical factor is, the more frequently [T3-T1] is used, and the less frequently [T3-T3], [T2-T3], [T3-T2], and the other tonal patterns are used.

2. Improvements

My experiment can be improved in three ways. First of all, the participants: although all of my participants are native TM speakers, many of them are bilingual or even trilingual. This would have an impact on the results since they might have connections to mental lexicon influenced by other languages and dialects when performing the experiment. Second, the population of participants could be expanded to improve data accuracy: there are only 8 participants in total thus the results were not precise nor accurate enough. Last, there could be more data sets. There are only 19 sentences in total. If more sentences could be placed in the experiment and more data could be collected, the experimental error can be reduced and precision and accuracy of the experiment would increase.

VI. References

1. Hsu, Hui-Chuan. (2006). Zailun Hanyu de Shengdiao yu Zhongyin [Revisiting Tone and Prominence in Chinese]. Language and Linguistics 7.1. 109-137.

2. Yeh, Chia-Hsin. (2010). Comparison of Phonetic Naturalness between Rising-Falling and Falling-Rising Tonal Patterns in Taiwan Mandarin. Paper in proceedings of Speech Prosody. Fifth International Conference, Chicago, IL, USA.

3. Huang, Tzu-chuan. (2016). I-R tonal mapping: evidence from the 'cute talk' reduplication in Taiwan Mandarin. Paper in proceedings of the 28th North American Conference on Chinese Linguistics (NACCL-28). Brigham Young University, Provo, UT, the U.S. (To be published)

4. Zhang, Jie & Lai, Yuwen. (2007). Two aspects of productivity in Taiwanese Double Reduplication. Kansas Working Papers in Linguistics. 10.17161/KWPL.1808.1786.

5. Cherry, K. (March 14, 2022). How priming affects the psychology of memory. Retrieved from https://www.verywellmind.com/priming-and-the-psychology-of-memory-4173092

6. Bybee, Joan. (2002). Phonological evidence of exemplar storage of multiword sequences. Studies in Second Language Acquisition. 24. 215 - 221. 10.1017/S0272263102002061.