國立政治大學語言學研究所碩士論文

National Chengchi University
Graduate Institute of Linguistics
Master Thesis

指導教授：萬依萍 博士
Advisor: Dr. I-Ping Wan

中文顏色詞彙處理機制：

心理語言學的實證

The Lexical Process of Color Terms in Chinese: Evidence from Psycholinguistics

研究生：蕭裕台 撰
Student: Yu-tai Hsiao
中華民國一百年七月
July, 2011

The Lexical Process of Color Terms in Chinese: Evidence from Psycholinguistics

BY

Yu-tai Hsiao

A Thesis Submitted to the
Graduate Institute of Linguistics
In Partial Fulfillment of the
Requirements for the Degree of
Master of Arts

July, 2011

ACKNOWLEDGEMENTS

有言，有我

「語」，從言吾聲。細解其部，有言有我。

我不由得想起進入語言所的初衷。

將望遠鏡拉大至可以對焦過去 26 載的座標系，我發現自己的人生走向，是從幼時在電視節目的字幕上看到了「難」這個字。當時年幼不解其意，僅覺這字型真是詭異，小手在空中胡亂揮了一番，發現自己能看一眼便將這「難」字立就，心裡異之，成就感驅使著這小小孩長大。

抱持著對數科的無力感、英文的使命感，進入了中文系，勉強欺騙自己文學可以化解身為書呆子的格格不入。亦師亦父的許長謨教授的語言學課程開始讓我正視「語言」這個議題, 在潛意識中是多麼地被自己關注。導致我點起火種後，奮勇前往外文系再修了蔣炳榮老師語言學概論，以一種嚴肅的態度幫語言在自己青春中定義劑量。那時永遠不會曉得，蔣老師曾在成大教過一位以失語症專題在台灣一片語言學理論為宗的領域異軍突起的學生，後來在政大語言所擔任起所長一職，此青年學人即我的指導教授萬依萍老師。

緣份在舉起時難以言表的那一刻起，便以一種無法辨識的邏輯串聯起來。萬依萍老師提供了實驗室、失語症及客語等專案計畫，讓我能在其麾下參與執行實驗與台大醫院的語料採集，在計劃的參與中更接近我想要關注的議題，也讓這份論文的付梓埋下了契機。

說錯話，是我此生最害怕的事情之一；但我卻藉著語誤研究去叩問我最好奇的知識。

這份實驗性的語誤研究，在這條路上蒙受台灣許多學者的指教與垂顧，從 2010 年喜懷千金的指導老師催生下筆始，10 月在時任所長的黃瓊之教授與師大徐東伯老師共同指點下通過論文計畫書，為本文奠定良好的架構；11 月至中央研究院發表，接受語音學大師 Lieberman 教授及與會學者的評點；隔年 7 月在口

試委員徐東伯老師與政大林祐瑜老師的審查與口考下通過學位資格；復於同年10月在第十二屆NCL發表本篇論文相關成果，得以聆受同為研究Connectionism之成大心理系胡中帆教授的鼓勵與心理實驗設計上建議。

我覺得時間不夠用。就算時間以秒來計時，依舊不夠用來將此論文修得更臻完善。特別是經常受我「剝削」的指導教授，卻從不吝給我全方面的照護。

修業期間，有一些老師的影響舉足輕重。

在語言所迴廊上靠著電動輪椅滾動起大家思緒的蕭宇超老師，讓我碩一時能夠用自己的力量與夥伴赴泰國參加國際會議。每到心理的一個分岔口，都很痛苦，但是總在這危惴不安的時刻，能夠提供一頭栽下去的動力，看看那條路的盡頭是什麼風景。雖難於行，卻總是走得比我們想像中都長都遠。

曾經立下志向，每學期至少要修一門何萬順老師的課，再難的行路，一定要修練下去。句法學家就像是在山洞點起一盞燭火守夜的人，得耐得住寂寞，搖椅是唯一的娛樂，靈光乍現的當下，即便因驚訝而熄滅了滿室的光火，也在所不惜。總是為了領受這靈光乍現的感動而於洞口外積極回應。

更是無法忘懷詹惠珍老師在聽聞口試過關的當下，在停車場為我仰天長嘯的那一幕。若非詹老師的課，則無法牽起南島語和共處在同一座小島上的我之間的羈絆。

惠鈴助教關鍵時刻的擁抱，將我把傾斜的地標扶正；如炬般的目光關候著所上每個人兒的一顰一痛，彷彿身上任何一處的抽搐，都能讓你共感。因此看到你而能放心，蹲下拾起散落四處的行囊，繼續走下去。

博士班是一群深藏不露的學長姐匯萃之處，在朝來水溶溶的市集中裝成路人四處遊走，偶爾揭穿某個人的身分，便會迅速變出一朵炫目的筋斗雲，在嘖嘖稱奇的一刻笑著乘著離去。在口試的這一天，彷彿說好似的，不經意地出現在我必經的路上，慷慨的提供微笑和心靈力量，還有無私的信任。文山秋天的山林像一首詩，送著章和的楓香和泠泠的雨聲。

每天一大清早，在不覺晨昏的實驗室，進行無數個行政事務還有語誤誘發的實驗。電腦桌面總是開滿著大大小小的視窗，新的銀幕巴不得愈換愈大，一個人在比我房間寬敞的小小幽室中，於器官模型的注視下，臨著語音器材，埋首諸多文案與音檔之間。經常會有許多天使如佳珈、晉瑋和涵絜、亭伊在實驗室的各個角落相伴，關心著你的近況、關心著怎麼完成老師交派的任務、關心著近來的梗概和彼此的小未來。實驗室的窗外是一座供鳥禽和人遊憩的庭院，矗立著幾株蓊鬱的闊葉樹，偶爾開出幾朵美麗的花，謙虛的等人尋覓喑啞的鳥囀而發現。

人身上都有超乎尋常的能力，這個超能力的行使是自己無法意識、但他人能輕易感受的能量，讓疲弱之人得到支持。在梵文中，此為「加持」之意。身雖離心相近的宏瑋用上帝滿滿的愛提醒我早已被神認識、並關愛著，所以我的身邊有采君、孟英、怡璇、佩如這些好朋友不定期定量的投以關愛，在一個讓我很舒適的距離輕輕拍落肩上的塵埃，常常讓你們忍受我偽善的世故，但台北因有你們而有溫度。還有經常床邊閒談鴻志的室友孝從，相見亦無事的語言 97 人，不來常思君的中文 96 人，以及一同在師大林千哲老師的心理語言學相見恨晚的 Julia、Brandon 與 Grace。木柵的潮濕常常鏽蝕著如碳鋼般的意志，在這只有燭光的山洞裡，享受大夥身上各自帶著一點鏽蝕的傷痕幫彼此除去溼氣。

對雙子座的人來說，生活是一幅多邊形，擁有許多稜角，角度寬狹不一，但必能找到質量重心，平衡各個層面，並用一條線穿引懸掛於風中。婷婷老師用舞蹈訓練著我的平衡感，在平時說話慣了、研究慣了語言之外，教我用肢體說話。我們在教育及舞蹈這兩條平行線上相知相惜，與這群一同在救國團學舞的姐姐們、年輕一輩的小弟小妹們，燃燒著不知已過了幾載的青春。

對我來說，有一處最重要的碼頭名為葉赫，它是我身為人師的啟蒙地，若在我拋下船錨的當下，有個人會幫我用愛心與耐心、把繩索緊緊繫在碼頭邊的，那個人就是葉玉堂老師。念當初為了讓教學能更符合語言的真相與脈絡進入了語言所，他用一通電話提供了舞台讓我能實踐語言學中的種種思維，使學生以一種有

「理」可說的方式認識語言、學習語言。方寸大的教室，坐滿著一欠身就能挑起一袋米、跟你談經論道、逼急了瞳仁會流出怯怯神情的國中生，我們各自參與著追逐對方漸去漸遠的背影，告訴彼此：不必追！龍應台眼中的目送大概就是這種感覺了吧。

不知在多少萬籟俱靜的夜晚啃資料、撰寫，母親用堅定且擔心的口吻說：「好去睡了。」我總是用石頭般的冷漠回應。我以為只有母親可以了解我淡定之下塵土盡揚的雜緒，唯她能善解。睡不到五小時，一大清早坐著老爸的便車、或他作我的乘客前往木柵。每天短短二十分鐘的相處，從原本的不耐…漸漸了解他的一生。二位長久以來用青春為土、愛心為料的栽培與犧牲，只能以盡量不讓你們操心的自律與自強回報。甫出嫁的姊姊用她的方式一直關心著我這個做弟弟的；還有自胚胎時期比鄰而生的雙胞胎哥哥，這份論文的統計概念多虧有他提點，雖不至完善，但比起自個兒橫衝直撞到某個死胡同而不自知要好太多。一家人在一起，小時會貪婪的希望大家將時間停留在自己身上，我們渴望相愛、專情；等年事稍大，會希望時間加快些，開始擔心自己之於對方是否只是一廂情願；到了接近出社會的那一刻，會希望時間走得慢一些，把青春期邁出父母太多的步伐放輕放慢點，看清你們臉上歲月爬梳過的容貌。 對你們這一段情感真的很難寫盡，一直有股力量阻止我繼續延展下去，暫且用這一本論文表達對你們的感謝。

這本論文且戰且走，過了好長一段時間才勉強付梓。等待的過程很沉重，但這個重量卻將我自己愈貼近知識、愈貼近生命。在啟程前往菲律賓前夕，為輕如鴻毛的青春，留下一些痕跡。

這本論文沒有很偉大，但它很像我會做的事。

VITA

<u>Education</u>

2012/03	M.A. in Institute of Linguistics, National Chengchi University
2008/06	B.A. in Department of Chinese Literature, National Cheng Kung University

<u>Grants and Scholarships</u>

2009-2010	Graduate Student Scholarship National Chengchi University
2011	The Phi Tau Phi Scholastic Honor Society of The Republic of China

<u>Publication</u>

2011	"The Phonological Effect on Lexical Encoding in Chinese: Evidence from Stroop's Technique and Speech Error" 12th National Conference of Linguistics: Kaohsiung.
2010	“The Lexical Process of Color Terms in Chinese: Evidence from Psycholinguistics.” Seminar on Phonetics/ Phonology and the Evolution of Language, Institute of Linguistics, Academia Sinica: Taipei.

2009 "A Phonological Analysis of Disyllabic Mandarin Onomatopoeia.", with W.-C. Yep, M.-T., Chang, 42th International Conference on Sino-Tibetan Language and Linguistics, Payap University: Thai.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

國　立　政　治　大　學　研　究　所　碩　士　論　文　提　要

研究所別：語言學研究所

論文名稱：中文顏色詞彙處理機制：心理語言學的實證

指導教授：萬依萍 博士

研 究 生：蕭裕台

論文提要內容：(共一冊，字，分五章，並扼要說明內容，共 741 字)

長久以來，序列式詞彙處理模式以及互動式詞彙處理模式經常爭論孰為合理的語誤生成模型。這份研究企圖利用心理學實證角度探究此議題。原屬於心理學範疇的史楚普技術(1935)或可提供語言學多方面的視角，尤其在我們處理顏色詞彙時詞彙譯碼歷程上與視覺表徵之間的關係。因此我們試圖將此經典心理學技術轉化成宜於語言學研究的語誤誘發實驗。藉由控制變項間的語音近似度，創制了四項實驗，分別是：顏色唸名實驗、顏色誦讀實驗、史楚普唸名實驗、同音詞唸名實驗。結果顯示，即便受試者對語音近似度並未在單項實驗中產生語誤數量與反應時間的差異，但它在不同實驗任務之間卻誘使受試者產生不一樣的反應結果，包括數量上及時間上的顯著差異。

此外，我們針對語誤，分析其與目標字之間的關係，研判部分音韻效應是否對詞彙處理網絡造成顯著影響。結果顯示，史楚普效應、音韻結構、音法限制以及聲調效應在語誤數量中均出現顯著的生成量；音節首、韻部以及母音則並未出現顯著效應。史楚普技術除了在語誤證據中為文字閱讀提供自動化效應的解釋外，也讓我們看到音韻效應在不同視覺任務指派中產生顯著差異的結果。如此看來，互動式模型提供了較為簡潔的解釋。來自二元視覺刺激(視覺色彩與顏色詞)的詞彙競爭，該框架可以合理提供理論基礎，並解釋不同視覺任務之間音韻關係依存度的不同。

另一方面，我們亦援引同樣技術設計出以音韻單位為導向的唸名實驗，討論語言輸出前規劃單位之議題。我們發現，僅含聲調的音節以及無聲調的音節可在詞彙網絡中做為語言規劃之心理處理單位，促進詞彙處理效能。其他單位如音節首、介音、母音、韻尾、聲調則無法出現顯著效應。同時亦於實驗中發現，中文聲調於規劃階段時，應是屬於詞彙結構上的聲調，而非純粹的聲韻調。

關鍵字：詞彙處理模型、史楚普效應、音韻近似度、語誤、語言規劃單位

Abstract

Serial-ordering model and interactive processing model have long been discussed as whether people process languages in a sequential level of processing or the consequence of levels interacting altogether by looking at speech errors. Stroop technique (1935) in psychology could give linguistics some insights on the relation between lexical encoding and visual representation when people process colors. We tried to adapt this classical psychological experiment for an experiment of speech error elicitation. By means of controlling the phonological similarity, we created four experiments: color naming, color reading, Stroop naming, and homophonous naming tasks. The result first showed that even though phonological similarity did not induce significant difference in error amount and response time within single task, it still caused significant difference in these results among these tasks.

Second, after analyzing the linguistic relation between targets and speech errors, we found that Stroop effect, syllable structure, phonotactic regularity, and tone induced apparent effects in error generation, while initial, rhyme, and vowel did not. Stroop effect not only provided the evidence from speech errors where character representation was an autonomous mechanism in lexical process, but also provided a fact that phonological effect would impacted differently on amount of error according to the type of visual task. It seems that interactive account could help explain the result easily, as the competition from dual visual inputs could be given a theoretical basis to account for the phonological dependency according to certain visual task which subjects were assigned to.

Finally, we also discussed the issue of advanced planning unit in lexical process. In the shared unit task with Stroop technique, we found that only the units of tonal syllable and bare syllable could serve as possible planning units in lexical network, and tone in Chinese should be attributed to a type of lexical tone in planning, rather than a pure phonological tone. To sum up, the purpose of the study attempts to provide empirical evidence to examine the above issues.

Key words: Lexical model, Stroop effect, phonological similarity, speech errors, advanced planning unit

Chapter 1

Introduction

The operation of language is the composition of lexicons and mental rules constructed in mind. The way in which the basic units are constructed could reflect the structure of an utterance, even though it is a process of interaction from several linguistic components: syntactic, semantic, and phonological layers etc. These layers could be attributed to human's linguistic competence that conveys the information to listeners and to process the linguistic signal from context. In such a language system, it doesn't always lead to a perfect and well-constructed linguistic form in daily language use.

With the evidence of speech error, the inconsistence between linguistic competence and language performance exists. Speech error has been an important source to study the linguistic mechanism and organization of cognitive ability in linguistics and psycholinguistics because any tiny problem (or noise), which happens to the process of language, will result in slips of tongue (Stemberger, 1984; Martin et al., 1996). Speech errors always reflect the failure or malfunction of their corresponding mechanism and linguistic structure (Fromkin, 1973; Shattuck-Hufnagel, 1979).

In addition to errors in naturalistic settings, psycholinguistic experiments have been conducted to induce observable speech errors and apply the data to simulate the processing model (McClelland & Rumelhart, 1981; Dell & O'Seaghdha, 1992; Sevald et al., 1995; Dell & Juliano, 1996; Roelofs, 1996; Dell et al., 1997; Levelt et al., 1999; Dell et al., 2004; Martin et al., 2004). On the other hand, several linguists took the corpus of speech errors to elaborate the linguistic structure and visualize model of speech production (Fromkin, 1973; Shattuck-Hufnagel, 1979; Garrett, 1980 & 1988; Dell & Reich, 1981; Stemberger, 1985; Martin et al., 1996; Levelt et al., 1999; Wan, 1999 & 2007). As a representative of serial model, Garrett (1980) proposed that the linguistic levels should be independent, and linguistic signals are processed step by step. However, as to an interactive view of language mechanism, not only Dell and Reich (1981) had observed the phonological similarities in word substitution, but Stemberger (1985) assumed these layers should be interactive when accessing lexicon, and the feedback from the form (phonological) representation could cause the wrong activation back to the lemma level to retrieve the wrong lexicon which is not intended. Sometimes, speech error might come from more than one layer because of shared linguistic information between target words and error words, such as phonological-semantic error. The two main models of language process, serial and interactive models, always adopt speech errors elicited from experiment or corpus observation to help them simulate and infer the possible route and operation of lexical process.

In this study, color will be the main focus of lexical process when we deal with the issue on lexical process. The semantic term 'color' will be specified throughout the study. Color is a universal concept, which maps to a language-specific lexical term according to languages. Color refers to a taxonomic relation in semantic field. For example, as shown in figure 1-1:

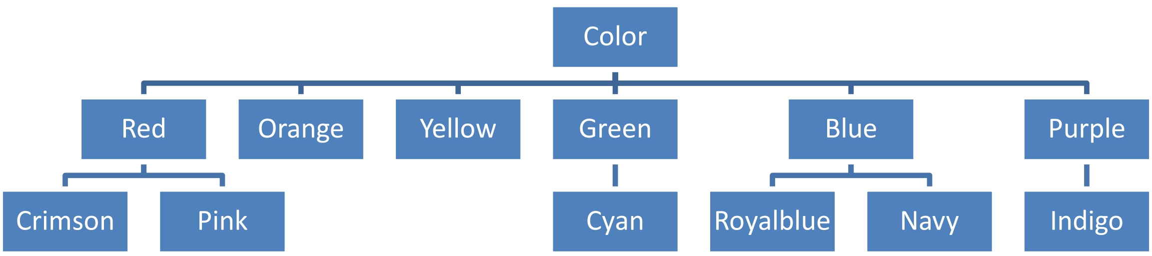


Figure 1-1. Color Taxonomy in English

In figure 1-1, the semantic hierarchy tells that the colors can be subcategorized into detailed ones, as crimson and pink of red, indigo of purple. Apparently, in the semantic field, these color terms are structured in a family which is superordinated by a general concept: color. With the view of semantic network, the hyponyms, dominated by "color," are connected tightly. In this study, we would like to adopt color to test whether the naming or reading tasks would produce speech errors inside or outside of the semantic field, and whether phonological similarity would induce speech errors during lexical processing. With regard to human's visual competence, Stroop's technique (1935) could be a classical experiment to study the speed of naming and reading. Besides the terms, the ink color of the term can also convey a semantic concept through visual channel. If we create the inconsistence between the color term and the visual color of carrier, the competition of the semantic and phonological information would induce some kind of "noises" during the lexical activation. The chance of speech error, or we can say the output through incorrect activation, might get increased in speaking. Stroop (1935) conducted a series of psychological experiments to test the reaction times among the color naming task, color term reading task, and the task of naming the visual word with different colors. It suggested that the reaction time for naming would be increased and prolonged by the presence of conflicting visual concepts, word and color at the same time. He found that the interference effect occurred when human processed the linguistic unit with visual, physical or linguistic conflict at the same time, which is well-known as Stroop effect.

In this study, we would like to conduct a series of experiments to extend Stroop's work. We adapted the classical experiment for an ideal one for Chinese, and we observed and analyzed the speech errors by means of controlling phonological similarity. We hope to understand the status of phonological similarity in lexical process. There were five experiments: color naming test, color reading test, Stroop naming test, homophonous naming test, and shared unit test in this study.

The organization of this study includes: literature review will be introduced in chapter two, focusing on the mechanism of naming and reading tasks, production models, planning unit and linguistic effects. The proposals will also be outlined in the same chapter. Chapter three will elaborate the stimuli organization and procedural design for each experiment, including the color naming and reading task with or without color competition. Chapter four will present the data on reaction time among experiments and speech error analysis so far. In this study, we try to figure out: through the induced speech errors, we would like to analyze the phonological relation with their targets and discuss linguistic effects between target and errors. Based on the amount of each error type and the reaction time, we would like to compare which lexical processing model might be more appropriate to simulate the process of lexical activation in Chinese which has been proposed by many psycholinguists.

Chapter 2

Literature Review

Speech errors and response span will be mainly concerned in this study. The lexical process models will be reviewed in the first section. We will examine whether the structure of speech error and speed of lexical process would induce linguistic effects, which will be reviewed in section 2. Besides, the issue on advanced planning units in language production will be reviewed in section 3. The interaction between sound activation and lexicon receiving, linguistic effects and the advanced planning units will be the main issues of processing model in this study.

2.1 Models of Speech Production

Lexical access models have been used to explain mechanism in lexicon storage when we process language. There are two main debates on how cognition system processes the lexicon information: serial-ordering model and parallel interaction model. The diverse routes in lexicon processing induce dissimilar patterns of sound and meaning process when perceiving as well as encoding lexicon. However, the assumption in any of the models should attain the sub-goal—simplicity for cognition process.

In this study, we also put focus on the interaction between sound activation and the lexicon retrieving. The selection of semantic concepts and phonological information will be mainly concerned in the following section. In addition, the goal of this research is to compare the serial model to the interactive model, and testify which side could have proper explanation on the lexical production.

2.1.1 Utterance Generator Model of Speech Production

Fromkin (1971) attempted to account for various types of speech error, and the

Utterance Generator Model was presented in a top-down manner without any feedback loops in language processing. Fromkin (1986) proposed that speech error could give more insights on the organization and representation of lexical items. The Utterance Generator model (Fromkin, 1971) is presented in Figure 2-1:

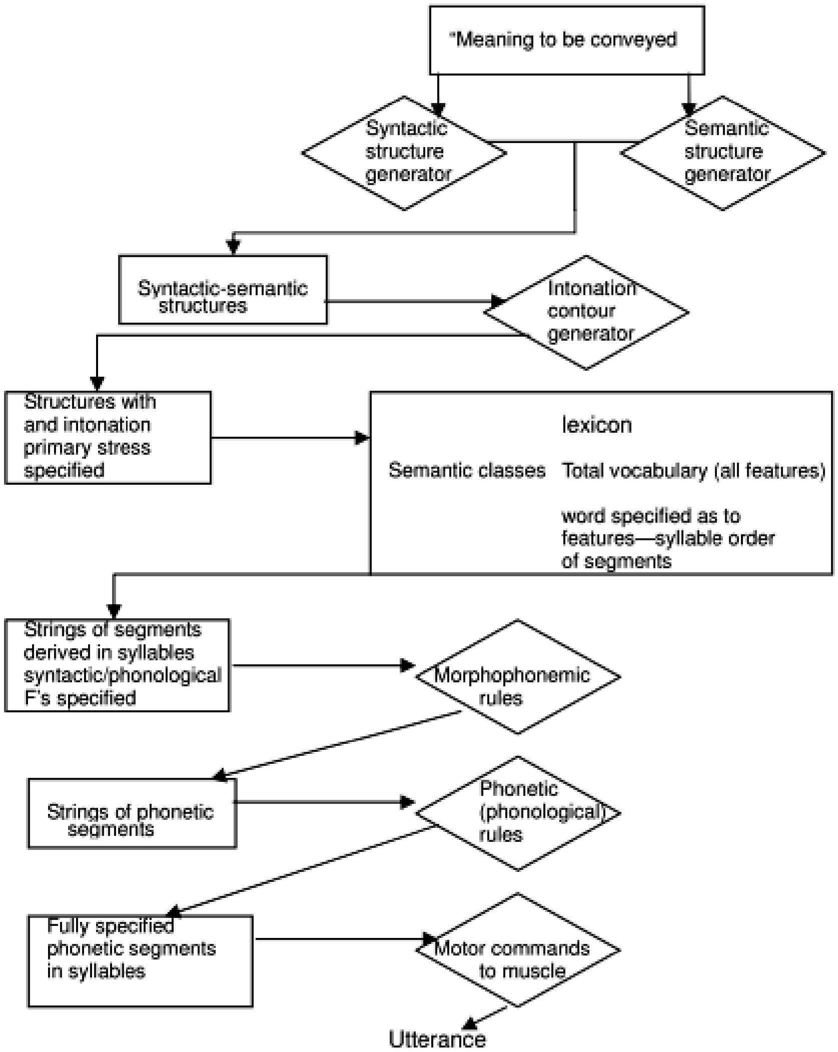


Figure 2-1. Fromkin's Utterance Generator (from Fromkin, 1971)

In Fromkin's model, the process of lexicon is divided into five stages: syntactic structure, semantic features, phrasal and prosodic assignment, phonological rules, and motor command. In Figure 2-1, the rectangular boxes in the diagram stand for the linguistic representations at each level; the diamonds symbolize the processes that translate each level of representations into the following one. The stages, representations, and processes are outlined below:

Stage I. A meaning to be conveyed is generated.

Fromkin assumed the model could allow the generation of more than one message at stage one. Fromkin quoted Butterworth (1980) and supposed that an adequate model of speech production should account for "competing plans" at all levels. This leads to one kind of speech errors at this stage—syntactic blending, such as *How long **does** that **has** to…have to simmer*. This model allows multi-representations at any processing levels.

Stage II. The message is mapped onto a syntactic structure.

After the semantic features and organization are outlined at first stage, a syntactic framework of the message is created, and semantic information will be mapped onto these grammatical structures. At this stage, a semantic-syntactic structure is marked as lexical nodes in the phrase marker for selecting words or morphemes from the lexicon at next stage. The grammatical category and syntactic structure will be specified at this stage.

Stage III. Intonation contours are generated on the basis of the syntactic representations.

Fromkin supposed sentence and phrasal stress must be assigned before lexical selection occurs. Syntactic structure always determined the primary stress and intonation contours, and Fromkin assumed this prosodic assignment should be independent of any other level.

Stage IV. Words are selected from lexicon.

Now the message is represented as a syntactic structure, with semantic, syntactic features, and prosodic information well-specified. The lexical items are not fully specified. It means that the affixes are not spelled out phonologically, and the stems are just at a stage of phonological features being specified as well as the syllable positions serially arranged. At this lexicon stage, semantically or phonologically similar words might be selected instead of the intended word. Fromkin also pointed out that mapping these words onto the grammatical structure may cause phonological segments or features to be dislodged out of their original sequence.

Stage V. Phonological specification

After the selection of lexical entries, the phonetic segments are fully specified in the specified forms and the phonological rules operate at this stage.

Stage VI. Generation of the motor commands for speech

At the last stage, the phonetic features bundles of segments or full syllables are encoded onto motor system for the following series of articulation.

This model advocates that the unit which induces speech error can be a segment or a feature when mistakes happen at later stage (stage IV and V). Error occurring implies the malfunction in one of the linguistic operators. However, Fromkin & Ratner (1998) commented that this model did not provide enough details why the major categories (nouns, verbs, adjectives) seldom shifted or altered in speech error, while functional categories and inflectional morphemes are. In addition, as Butterworth (1982) pointed out, if the model allows more than one sentences to be generated at initial stage, we might need more postulation of mechanisms (or rules) to generate the numerous proliferation of representations at the following levels. These postulations might make this generator model "more expensive" in explaining the speech generation.

2.1.2 Levelt's Lexical Model (1989, 1999)

As to the model of serial access, Levelt's model is also a unidirectional operation in language processing, as shown in figure 2-2.

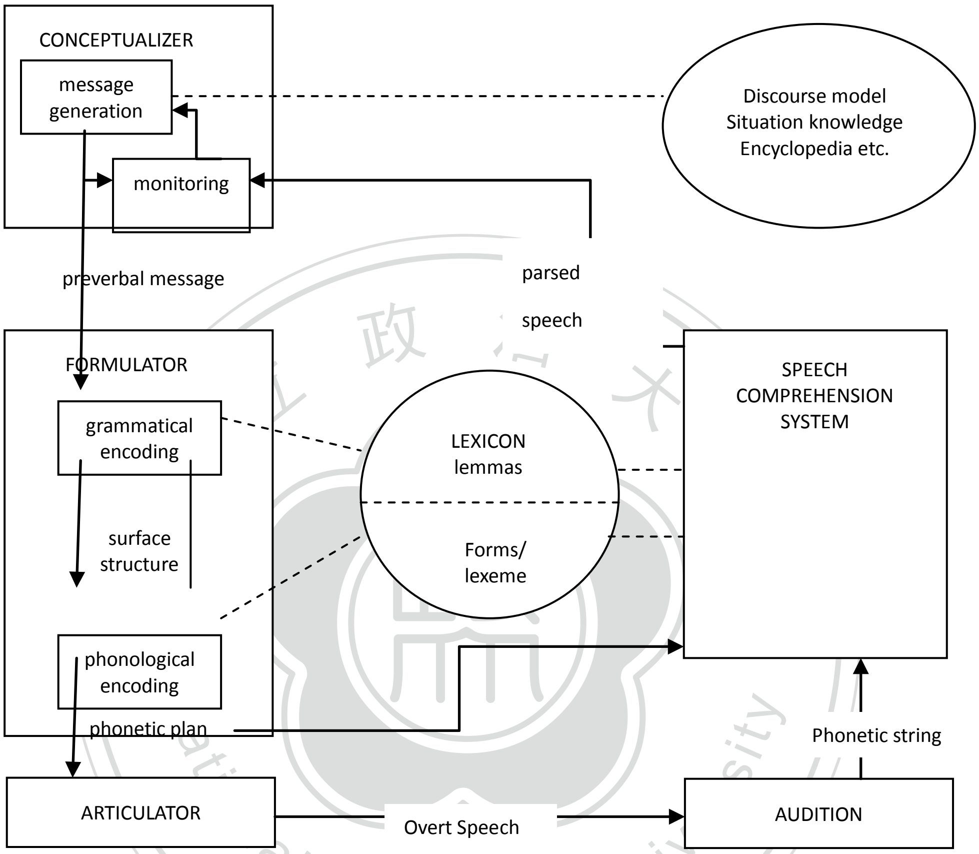


Figure 2-2. A blueprint for the speaker. (Levelt, 1989).

Levelt proposed that from the concept formatted to the sound activation, the word activation is step-by-step at each procedure in a rapid manner. Two levels are distinguished during the orderly processing: lemma level and lexeme (form) level, which are both linguistic functions in formulator. The former is a semantic-grammatical encoder, where concepts are passed on and assign proper grammatical function to each concept; the lexeme level refers to a phase which takes syntactic framework to generate phonological units (syllable and segment). Levelt believed that the two encoders are stored and operated separately. Once a category is selected, the grammatical encoder (lemma level) would produce appropriately ordered sequence for lemmas.

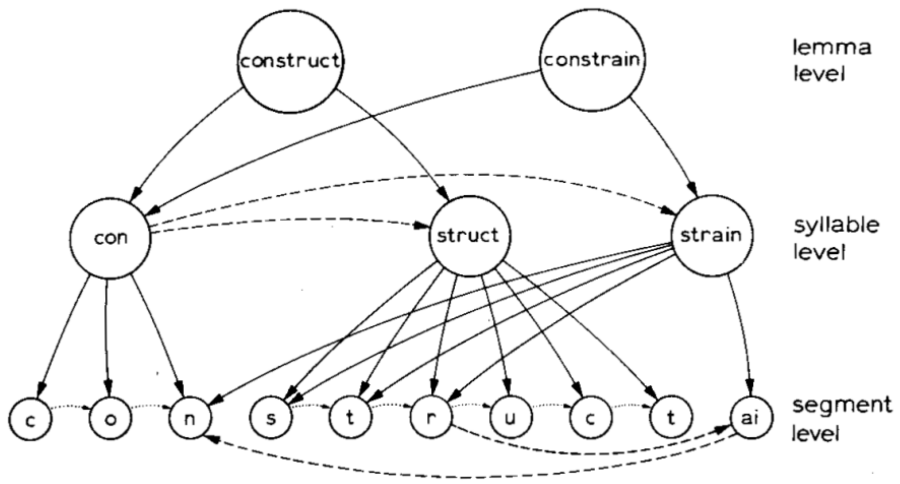


Figure 2-3. Example of activation-spreading network. (Levelt, 1989)

Figure 2-3 appears the phonological-segment access in detail. The phonological encoder (lexeme level) then adopts the grammatical outline that is specified at lemma phase and generates the phonological plan for the utterance. The lemma information is retrieved before lexemic information. The outline of these levels in Figure 2-2 listed below:

Phase I. The Lexical Level (or Concept Level)

The image activates the lexicon module carrying all the information the brain has stored about "construct." Each node is distributed in the network of connected neurons in brain. Adjacent lexical nodes exists semantic relations are also activated.

Phase II. The Lemma Level (or Category Level)

The activated concepts are passed on to this level, where the syntactic information is well assigned to each lemma. The grammatical features include the word order, case marking, thematic structure, and so forth, which are specified at this level. Lemma is semantic-syntactically specified without segments filled in the position. If there are more than one activated nodes, usually the highly activated node wins. But the more competing concepts interfere, and the longer it takes to generate the intended lexicon.

Phase III. The Lexeme Level (or Form Level)

After the intended concepts are syntactically specified at the lemma level, all the phonological information starts to be specified and accomplished at this level. Tip-of-the-tongue always occurs at this level, when a given lemma was not sufficiently activated to have completed phonological encoding at lexeme level.

With above "blueprint" of speech, Levelt further explained that self-monitoring effect operates and helps speaker detect the utterance with error; in addition, this operation may also makes speaker generate non-word output. It could help explain the lexical bias effect in phonological errors rather than phonotactic errors.

The main difference between serial-search model and the following interactive activation model is whether the linguistic levels could be interactive. Some speech errors, such as mixed errors, represent the interaction of sound and meaning information, which are difficult to judge whether semantic level or phonological level encodes wrong first. If the amount of semantic-phonological error (mixed error) goes salient, it could be better explained that interactive fashion could help explain error production further than in serial model.

2.1.3. Interactive Activation Model (Stemberger, 1985)

Apart from the former serial-ordering model, which is a one-way process in language production, there are some models that provide different viewpoints. Stemberger (1985) proposed that several levels are accessed in interactive fashion for the goals of the greatest use of finite properties in cognition system and the most minimal cost in process. The general idea could be symbolized in figure 2-4 below:

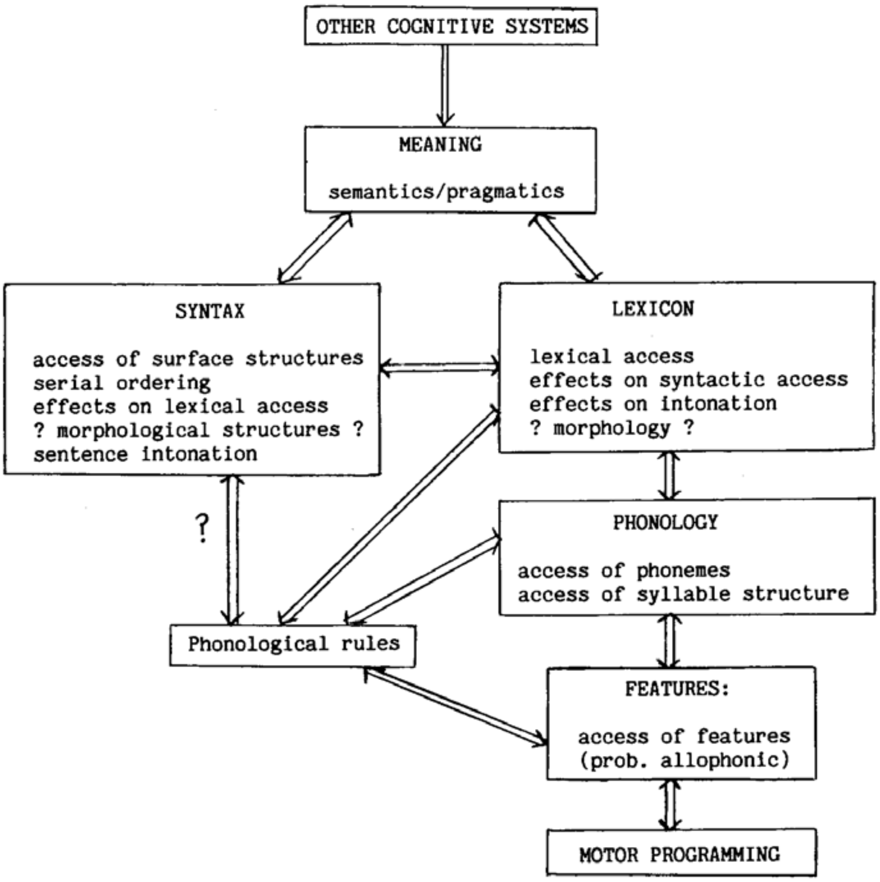


Figure 2-4. Interactive Activation Model (Stemberger, 1985).

There are two critical elements in an Interactive activation model of cognitive system: units and links. The basic driving force is called activation, a measure of the activity of a given unit. Highly activated units have large effects on other units; whereas the less activated units do not. The levels, involving segmental and feature levels, phonological and lexical levels, and lexical and syntactical levels, are supposed to be interactive in processing. In general, links between levels tend to be activated, which might induce some other influence on adjacent levels during the lexical access. Stemberger supposed that avoiding the interaction of levels in language production might cause greater complexity in lexical process. With several linguistic levels, simplicity argues for the interactive activation of components. The following shows the interactive nature of language processing in Stemberger's activation model:

1. The interaction of word and syntax

   The existence of particular lexical item often influences the phrase structure which is accessed. For example, the existence of an adjective might lead to the activation and selection of an NP with an adjective node. The mapping between lexicons and argument structure can be specified and post-checked between the levels.

2. The interaction of lexical items and phonological information

   There are errors where one word substitutes for another not because of a semantic relationship but because of similarity in phonological form, (ex.) malapropism[1]. As a lexical item being activated, it passes activation onto its subsequent phonological units. As the segments, syllable patterns, and stress patterns of the word become activated, they also spread the activation back to the lexical level as feedback, not only reinforcing the activation of the activated route and nodes, but spreading secondary activation to all phonologically similar words.

3. The interaction of the segment and feature levels

   The errors with pure phonological relation arise also due to feedback from the feature level to the segmental level. Feature errors involve the apparent misordering of a feature rather than an entire phoneme. Stemberger suggested that competing phonemes are given activation through feedback from the feature level, and a secondarily activated phoneme causes the syllable structure to be modified, and leading to allow a new phoneme

[1] Malapropism refers to one kind of speech error which exists phonological similarity, but with less semantic relationship. For example, "Texas has a lot of electrical votes." The word 'electrical' is substituted for the intended 'electoral.'

substituted.

On account of the various error types in speech, Stemberger supposed that the "noise" among the levels always leads to diverse errors during the cascading and feedback activation, such as phonological, semantic, or inflectional types, especially when we want to find a processing route to generate the production of mixed errors with more simplicity.

2.1.4. Connectionist Model (Dell, 1986, 1992)

Dell's (1986) lexical process model of speech production is known as a connectionist model. The different point from serial-ordering model is that connectionist model allows the activation weight sent back to the former nodes which have been processed. It is depicted as the "feedback" of activation among the levels interacting with each other. The general process of the connectionist model is depicted in figure 2-5 below as well as the diverse representation levels dealing with external signals in figure 2-6:

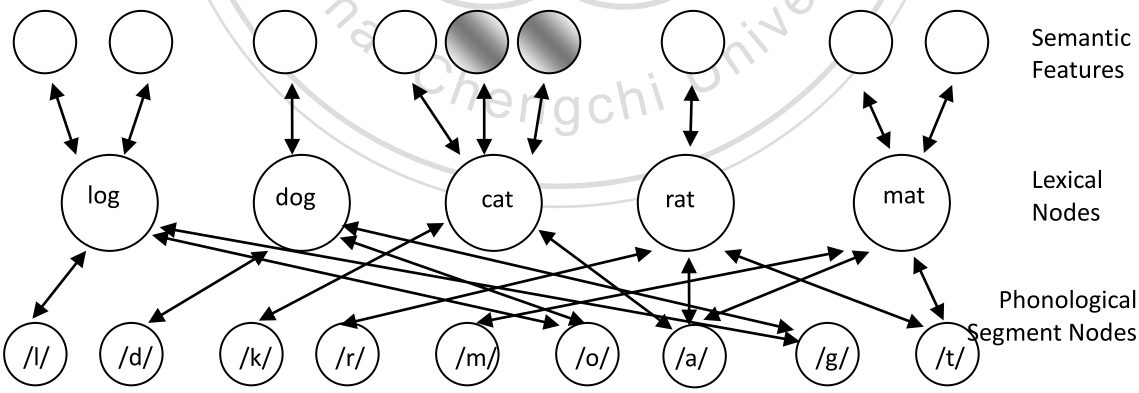


Figure 2-5. Lexical network in the spreading activation production model.

Like discrete stage models, non-discrete spreading activation models distinguish semantic-conceptual units, lemma units, and form units. The units in this model are organized into a network in which the connections allow for a bidirectional spreading of activation between units at adjacent levels. For example, the lexical word unit can pass activation downward to phonological level and backward to semantic level. As supposed in Dell and O'Seaghdha's work (1991), lexical access involves the following six steps:

(1) The semantic units receive external inputs.

(2) Activation spreads throughout the lexical network in a non-discrete fashion.

(3) The unit which receives the most activation is selected.

(4) When the word unit is ready for phonological encoding, it is given a triggering jolt of activation, whether it's a multi-word utterance or a single-word utterance.

(5) Activation continues to spread in lexical network, and the appropriate phonological units become significantly activated.

(6) The most activated phonological units are selected and linked backward to the slots of the word-shape frame (morphological structure) to implement sounding on those words, and so as to the frame of syntactical level.

Dell argued that the interactive model lies in its ability to account for a range of speech errors, particularly the effects of the similarity of the target and error in utterances. For example, Fay & Cutler (1977) proposed the fact that the all substituting and replaced words are very often from the same grammatical class, and it has been regarded as evidence that these substitutions are lexical errors rather than segmental errors. However, Dell believed interactive perspective among linguistic levels provides better explanation for those form-related (phonological-related) errors, which also imply the influence from grammatical and semantic levels.

Phonological word substitution can be attributed to a target being replaced by a phonologically related word, not just segments exchange. For example, as in figure 2-5, the word *cat* is activated, and, in turn, /k/, /æ/, and /t/ are appropriate units to be activated. Meanwhile, these active phonological nodes pass activation back to the lexical level to activate the word nodes relevant to them, such as *mat* or *rat*. Once the *mat* gets the most activation, the pure phonological error will be induced (for the rhyme sharing in phonological units); once *rat* gets the most activation, mixed error (with both phonological and semantic relation) will be generated. Therefore, mixed errors could be attributed to an interactive influence from both semantics and phonology in lexical access, which could make the whole model explain mixed errors not as costly as other discrete models do.

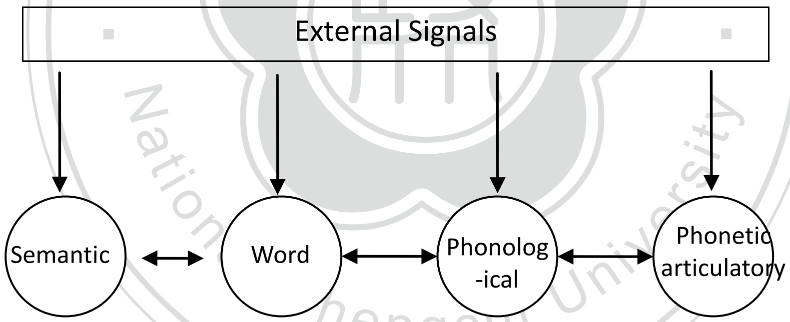


Figure 2-6. A simplified network in which external signals are received in sequence by nodes at four representational levels (Dell & O'Seaghdha, 1991)

Figure 2-6 shows a simplified network consisting of a single column of semantic, word, phonological, and phonetic nodes, and the different representations can also be interactive to induce the influence on the accessing efficiency.

Adapted from Levelt's serial model, there are also linguistic levels for nodes to process information from diverse sources, which might be from concept, word, sound levels or other linguistic layers. The difference is that Levelt's model is unidirectional-oriented, while Dell's is bi-directional interaction (stemming from the tradition of Stemberger, 1985). As to the main concern of mixed speech errors in this study, interactive processing model could give speech errors better explanation on many sound-meaning errors because the incorrect activation could occur in more than one layer during lexical process. In current study, we assumed that connectionist model might have better account for the lexical access and the speech errors in a wider view, but we need more empirical evidence from the experiments. On the other hand, if concepts competition (from literal and visual information) occurs in lexical process, whether the mixed errors would be induced more is mainly concerned in this study.

2.2 Linguistic Effect in Previous Studies

In order to make an objective judgment on the lexical access model that could well-account the speech errors, we have to know the internal structure among these errors, especially for the linguistic effects on the semantic-lexical level, lexical-phonological level, and phonological-phonetic level between the target words and errors. In this section, some linguistic effects will be reviewed, such as initialness effect, phonological similarity effect, rhyme effect, syllable structure effect, tone effect, phonotactic regularity effect, and Stroop effect (originally used as psycholinguistic term).

2.2.1 Initialness Effect

Most syllables have an initial (onset). Some languages restrict onsets to be only a single consonant, while others allow multi-consonant onsets according to various language specific rules. Chinese is the language allows both null and occupied initial in syllable. However, the status of prenucleus glide, attributed to initial, rime or independence, still keeps undetermined in Chinese. One thing is for sure, initial consonants are more likely to slip than non-initial ones (Dell & Juliano & Govindjee 1993). Take an initial metathesis error as example:

(1) tonal phonology → fonal phonology

The initial [t] was substituted by the initial [f] of the following syllable in an anticipatory manner. Initials tend to slip more than other phonological structures, which can also be seen in the study found in Boomer & Laver (1968), Mackay (1970), Dell & Reich (1981), Dell (1986), Shattuck-Hufnagel (1986), Stemberger & Treiman (1986), Stemberger (1989), and Dell & Juliano & Govindjee (1993).

Dell (1986) indicated that the initial consonants may activate stronger connection weight from higher levels and thus tend to induce slips. Dell et al. (1993) noticed that there was approximately 80% of consonant movement error involving syllable onset, such as the error in (1), while there was still 40-50% for non-movement errors. The further point they proposed as typical explanation for initialness effect has been that initial consonants of words are structurally salient and distinct in the phonological frame. The frame corresponds to Shattuck-Hufnagel's (1987) word onset and MacKay's (1972) syllable onset, which are more detachable than the other structures of a syllable. This effect are assumed by Dell et al (1993) to reveal the manner of phonological structure as they are used in the retrieving the sounds of words.

2.2.2 Rhyme Effect

Rhyme effect intends to examine if interacting units of lexical errors share their rhymes, including CV, VC and CVC. It reveals the hierarchical structure of syllables. According to Syllable Structure Hypothesis (Hockett, 1967), a syllable could be divided into two subgroups—consonant clusters formed one while the vowel and the final consonant was the other. Thus VC sequences were frequently preserved in phonological errors in previous researches (e.g. Nooteboom 1969, Mackay 1970, Shattuck-Hufnagel 1979, Rapp & Samuel 2002).

In order to test whether the hypothesis is also applicable in Mandarin, two types need to be approved— a) onsets should be more likely to err than codas, and b) VC sequences, as a unit, tend to induce more errors than CV sequences do. If, in this study, the onset induces less errors or VC intents to induce more speech errors, there could be a rhyme effect.

2.2.3 Phonological Similarity Effect

Similarity effect means that phonetically-similar segments tend to interchange with one another. (e.g. Nooteboom 1969, Mackay 1970, Fromkin 1971, Fay & Cutler 1977, Shattuck-Hufnagel 1979, Levitt & Healy 1985, Stemberger 1985 & 1989, Wan 2007). Fay and Cutler (1977) asserted that mental lexicon was phonologically-arranged based on a distinctive feature system. During the process of lexical selection, the phonologically-similar item in the neighborhood of the intended unit is apt to be selected and thus generated lexical errors. It can account for why the exchanging error in (2) occurred.

(2) reading list → leading rist

The two initial consonants [l] and [r] share many except for one phonological features: [±lateral]. Fay and Cutler (1977) observed 156 malapropisms and found nearly 25 percent differed in only one feature. Therefore, in the process of word activation, the significance of phonological similarity effect illustrates the cognitive status as segments. Wan (1999) examined speech errors and found feature values can alter or be substituted. Wan agreed with Fromkin's study in which features have cognitive status in speech production, but only the segment has function of segments primary planning (Fromkin 1979)[2].

2.2.4 Syllable Structure Effect

Syllable Structure contains two levels of meaning: syllabic constituent interacting and syllable pattern interacting. The first level indicates the errors within syllable, especially single-segment substitution. It is also called Syllabic Similarity Phenomenon, which indicates that interacting segments occupy the same syllabic constituent in the syllable, i.e. onset substitutes for onset, nucleus replaces nucleus and coda substitutes for codas. Mackay (1970) and Wan (2007) found that such kind of replacement errors occupy 98% and 99.22% respectively. This phenomenon can also be seen in what Boomer & Laver (1968), Fromkin (1971), and Shattuck-Hufnagel (1979) have found.

In addition, the second level indicates the interacting syllable patterns are usually the same among substitution errors, i.e. the CVC pattern of errors interact with the CVC target, and CV pattern interacts with the target in CV pattern. Such similarity of syllable structures between interacting words is described in Dell's (1988) model, as shown in figure 2-7.

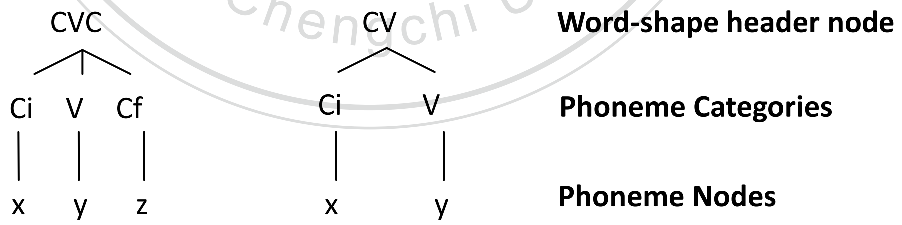


Figure 2-7. Syllable Constituent and Structural Hierarchy (Dell, 1988)

In Dell's proposal, each word node in the lexical network connects to a word shape 'header' node with the CVC or CV pattern, and it links to Phoneme Category to

[2] Psycholinguists proposed that segment could be an advanced unit for planning, such as Fromkin (1971). Current phonologists, based on OT approach, supposed that phonological feature should be a basic unit in lexical process.

specify the syllable constituent, i.e. Ci (initial/ onset), V and Cf (final/ coda). Finally, each category node connects further to all possible phonemes (/b/, /d/…). Under this framework, the number of activated phonemes is defined by the word-shape header, called ‘categorically trigged selection,’ which means the nodes with the same syllable pattern are activated, and the highest one will be selected in the end.

2.2.5 Tone Effect

The status of tone in Chinese has disputed whether the mental status of tone exists in lexicon or is pure phonological information.

With the view to tone independency, Packard (1986) observed that Chinese aphasics with a brain lesion in the left hemisphere experienced a deficit in producing tones that was qualitatively and quantitatively similar to the deficit in producing consonants. He argued that “tones, like consonants, are listed in the lexicon as unit phonemes … [and] the tonal phonemes, like segments, are composed of bundles of distinctive features” (p.220). Besides, Chen (1999) and Chen & Dell (2003) followed Packard and argued that tone is lexically distinctive, but tones were just like suprasegmentals in the underlying forms and became associated with rhymes or vowels at the phonetic level. They observed the tone status by means of implicit priming task. They found that the units of syllable-alone and tonal syllable would induce priming effects, while bare tone unit seldom plays the role of priming in implicit priming task. They proposed that the tone may be possible to be separable in lexical process, but it cannot be served as an independent unit to have priming or speech planning. Their explanation assumes that tones behave like segments in Mandarin Chinese because of the priming effect; they also assumed the tone to be less free because it alone induced little priming effect. Chen (1999) further stipulated that tones in Chinese were part of the phonological frame because it acts like the stress in

Dutch and English.

On the other hand, Wan & Jaeger (1998), Wan (2007) seek for the evidence form speech errors in Chinese and supported that tones are represented lexical underlyingly and ought to be part of the phonological organization of the lexicon. Wan & Jaeger (1998) observed the speech errors and pointed out that when one lexical item is substituted for another, its tone (level or contour tone) remains unitary; when its item is deleted, its tone is also deleted, further proposed that if two words have been blended into a single phonological form during lexical insertion, one of the two underlying tones, usually that of the selected rhyme, would be spoken in the errors. Besides, Wan (2007) and H. Chen (2008) disagree with J. Y. Chen's (1999) and J. Y. Chen & Dell's (2003) arguments, because the speech errors sharing the same tones with the targets outnumber those with the rhyme effect. (Wan & Jaeger: Tone effect 57.01%; Rhyme Effect 16.51%). The evidence showed that tone should not be associated with rhyme since the error rates of this pattern were not salient.

2.2.6 Phonotactic Regularity Effect

Phonological speech errors seldom create sound sequences that are illegitimate in language. Stemberger (1983) provided one of the rare counterexamples, as in (3):

(3) **d**orm → **dl**orm

Error (3) showed that this neuroticism error is a violation of phonological constraint of English prohibiting /dl/ initials, which is very rare in speech errors. By means of observation of error collection, Meringer and Mayer (1895) initially note this effect, and it was termed the "first law" of speech errors by Wells (1951). Stemberger (1983) noticed there were less than 1% phonological errors which were violations of phonotactic constraint. Production theories claim much of the effect in terms of phonological frame. Dell (1993) indicated that frames are specified in such a manner which impossible sound sequences are proscribed. For example, in Dell's (1988) model, the phonological frame enforced the regularity effect by controlling which sequences of categorized slots were made available. It was thought that there is no available frame that will allow an illegitimate sequence, such as /dlorm/ in English, to be encoded in production. The studies of Dell, Stemberger, as former mentioned, and Fromkin (1971) all took this effect as solid evidence for the active use of phonological rules or frames in encoding phonological sequences of words.

2.2.7 Stroop Effect

This effect was named after John Ridley Stroop who first delivered the effect in English in 1935 in an article entitled *Studies of interference in serial verbal reactions*, and the Stroop effect refers to the color term reading overrides and is quicker than color naming, which was conducted by measuring the reaction time between visual representation and naming.

In psychology, the Stroop effect is a demonstration of the reaction time in naming task. When the name of color (e.g., "blue," "green," or "red") is printed in a color not refer to its name (e.g., the word "red" printed in blue ink instead of red), naming the color of the word is longer, and naming would result in more errors and hesitations than naming the stimulus which matches to its denoted color. Stroop's experiment included the following three main parts:

(1) The Effect of Interfering Color Stimuli upon Reading Names of Colors Serially: all the color term were printed in black to measure the effect of the black color as an interfering stimulus.

(2) The Effect of Interfering Word Stimuli upon Naming Colors Serially

The name of the color were appeared in different ink other than it referred to.

(3) The Effects of Practice upon Interference

The visual stimulus was appeared in squares of a given color.

The difference among the time span for reading the words printed in various colors, the same words printed in black, as well as the same colors printed in squares were measured to observe the interference of conflicting word stimuli upon naming colors. The result showed that the interference of conflicting word stimuli upon the time for naming colors caused an increase of reaction time for naming visual colors. In psychology, Stroop effect set up that the association process in naming visual colors is radically different from that in reading printed words. However, for linguistics, by means of Stroop technique, we would like to observe further whether the time course is different when the colors are well-controlled in sound similarity, and we could also analyze the phonological relationship between the target and the induced error. Phonological similarity, reaction time, and speech errors will be mainly concerned in experiment of current study.

2.3 The Planning Unit in Sound Production

The issue on advanced planning unit will also be reviewed in this section because speech errors can be evidence of the processing quality and precision from syntactic level to phonological level.

Research in speech production has long been focused on Indo-European languages. The WEAVER++ theory (Rowlof, 1997) asserts that the compositional morphemes of a word need to be spelt out before phonological encoding of the word can proceed. However, the postulation about the word encoding was not satisfied to stay valid in Chinese. Chen & Dell (2002, 2003) conducted an implicit priming task (Meyer 1990, 1991; Roelof & Meyer 1998), and noticed it seems not to be the case. Their results showed that: (1) syllable-alone unit can induce some priming with different tone, whereas tone alone cannot induce such effect; (2) "syllable + tone" (tonal syllable) is a possible unit in word production and orthography and the word's morpheme do not seem to play a role in language planning; (3) tone could be a metrical frame for the segmental syllable to attach to in Chinese; (4) syllable-onset cannot induce priming effect either, so it couldn't be regarded as a planning unit in underlying.

2.4 Literature Summary and Research Hypothesis

Researchers have investigated sounds activation in production mechanisms, as proposed by the former serial and interactive model, by means of error-inducing experiments (Dell 1986 & 1990 & 1992, Schwartz 1994, Chen et al. 2002, Martin et al. 2004, Schwartz 2006) and speech errors in a naturalistic setting (Fromkin 1971, Buckingham 1980, Stemberger 1984, Bloch 1986, Schwartz 1994). Both production models should be compared to identify the role of the sound, and how the elements are activated in these models.

Theoretical frameworks of serial-ordering and connectionist have debated on whether the error sounds are produced from a sequential level of processing or the consequence of levels interacting altogether. As to serial approach, the sound error was dealt with the interference within the chain of linguistic components, including semantic, syntactic, morphological or phonological generators (Fromkin 1971). In order to explain how the retrieved lexicon is proceeded to the next level to retrieve the sound, which influenced Levelt (1989 & 1999) to propose lemma and lexeme levels in explaining the phonological units are encoded falsely to the retrieve a similar lexicon (at lemma level), thus the errors are produced. Either target sound or error sound is uttered; serial-ordering model prohibits the possibility of backward activation in lexical processing.

On the other hand, interactive approach proposes bi-directional activation: not only spreading but also feedback during lexical activation. Speech error seems to be the noise between lexeme and lemma nodes, or we may say it is also a false connection (Stemberger 1985), on the way to the target lexicon which receives enough activation falsely. The phonetic output in the interactive mechanism is the product of mutual interaction among syntactical, lexical and phonological levels, which in serial-ordering model is merely the product of single source, route, or Fromkin said the false encoding of morphophonemic rule results in error, or the Levelt said it is the false retrieving from lemma level to lexeme level. In Dell's connectionist model (1986), the output comes from the lexical-network organized in mind, also comprises semantic, lexical, and phonological layers, but is the co-effect of these levels. Sometimes speech errors result from the latter level spreading weight back to the former level. The feedback strength goes back from phonological layer to word layer and retrieves a lexicon framework with similar morphophonemic organization, spreads again to phonological layer, and then induces speech errors. The interactive simulation might imply the information of phonological units (feature, segment, or phonological constituent), pure semantic substitution, or the mixed error (interacting between phonological and semantic layers) about speech errors.

In current Stroop-technique study, there are some research hypothesis about the relation between sound activation and visual representation, which need to be dealt with in this study:

(1) Through the control of phonological similarity, we would like to know the trials of high phonological similarity would induce more speech errors than trials of low similarity, and see whether the controlled factor would cause difference of error amount among visual tasks. In addition, we should observe and compare the response time between the two groups. We would like to know whether phonological similarity serves as interfering or facilitation

effect in these tasks.

(2) Among speech errors, we plan to grade phonological similarity between target and error. We would like to examine which linguistic effects dominate in lexical encoding: initial, vowel, rhyme, tone, syllable structure, phonotactic regularity, and Stroop effects in this study.

(3) In order to look further into the issue of advanced planning unit during language production, carriers with diverse structures of shared units were recruited, such as the units of onset, vowel, rhyme, tone, syllable, tonal syllable. Response time and amount of speech error will be compared and discussed on their possibilities of serving as planning units.

(4) The status of tone will be concerned in shared unit test. If it is a phonological tone, it will induce facilitation in lexical process. If it is a lexical tone, the unit of tone should not induce any facilitation.

(5) According to the result of error amount and reaction time in the five tasks, where processing model could well account for the generated patterns. Serial and interactive accounts will be testified respectively.

This modified Stroop-technique experiment could provide supportive and empirical evidence to examine whether serial-ordering model or connectionist model could tell us more about sound activation during human lexical production. If we could take speech error, reaction time, and basic planning unit into consideration by means of controlling phonological similarity between linguistic forms (color term / carrier)and visual concepts (visual color), we could obtain more insights on how concept and sound are processed when human encode lexicon, especially there are dual inputs coming into visual channel.

Chapter 3

Methods

In this experiment, we modified the test items in Stoop's experiment technique (1935) and conducted experiments which we controlled the variety of phonological similarity and added homophonous materials into experiment. In the original work, Stroop designed the color term as visual stimulus with ink color other than the term indicated, such as the test item of "red" in color green, as reviewed in former chapter. Stroop technique was utilized into current study. In current study, all the stimuli were replaced with Chinese characters. In addition to the language variety, phonological similarity was an independent variable controlled in this study. The amount of speech error, types of speech error, and reaction time in each trial were dependent variables in the experiments. The amount of speech error and response time were served as judging facilitation or inhibition effect during color processing. Linguistic effects between target and error could help us observe whether there was any structural effect in lexical encoding. Phonological similarity was utilized to observe whether it would increase difficulty or error amount in lexical processing. Therefore, the criteria for phonological similarity, stimuli design, task procedure, and error analysis will be mainly concerned in this chapter.

There was a series of tasks which were tested for different purposes, which include: (a) Task 1: Color Naming Test, (b) Task 2: Color Reading Test, (c) Task 3: Stroop Naming Test, (d) Task 4: Homophonous Naming Test, and (e) Task 5: Shared Unit Test. The five tests were designed with different visual tasks. Subjects were asked to identify the colors they perceived or to read the words they saw, according to the conditions those tasks required. During task 1, subjects were presented a strip of color squares, and they needed to call the colors in order for each strip. In task 2, subjects had to read the color terms for each trial, and they didn't need to name the color of each term. Task 3 modified the Stroop's work (1935) into a Chinese edition. Subjects were required to name the ink color for each term sequentially. This was aimed to testify whether interfering effect of color stimuli upon reading color terms exists in Chinese, as that mentioned in Stroop's (1935). In present study, we created a list of homophonous stimuli corresponding to the color terms. Subjects had to name the colors for stimuli in task 4, as they did in task 3. Homophonous Naming Test used these colors' homophones to substitute for corresponding colors. As to the reason which we recruited in task 4, if homophone induces similar response time and error types, Stroop's effect might not be specialized to color terms, even though phonological similarity might have effect on both processes. If it induces various patterns of time span and error types, it appears that color terms might be enclosed as a semantic group at the lexicon process for further phonological encoding. If not, it implies that homophone serves the same role as color term in lexical process.

In task 5, each trial was replaced with other kind of characters, which shared one certain structure of phonological units with color terms. Subjects were trained to read and name the stimuli's color for each trial according to the tasks. Diverse shared units were supposed to induce interfering or facilitation during lexical process. Shared Unit Test was aimed to focus on the issue of advanced planning units among these color terms. We created a list of characters which shared different types of structure with color terms, such as initial, vowel, rhyme, bare syllable, bare tone, or whole syllable. The subjects' response time and induced speech errors among the four visual tasks were record for later analysis.

3.1 General Methods of Task 1 ~ 4

3.1.1 Subjects

We've recruited 22 college students (7 males and 15 females) whose native tongue was Taiwan Mandarin, aged 20-30 years old. All these recruited subjects took the four tests: naming test, reading test, Stroop's Task, and Homophonous Naming Task. In order to balance the practice effect and to avoid it inducing practice bias, half of the subjects were provided the order of color square naming, color term reading, Stroop's reading, and homophonous naming tests in order. The others were also provided the four tasks, but were presented as the order of homophonous naming task before Stroop's naming task. The ones with psychology, linguistics and art background were excluded from this study, especially the art major students who might be professional and sensitive to the subtle difference. When they identify the colors, they might probably generate many low-frequent color terms as their response, which might cause difficulty in data analysis. This study was aimed to recruit ordinary as well as normal speakers in color identification and to analyze the way they process concepts in color naming, instead of the way how they sort these colors sophistically and precisely.

Besides, in order to ensure that subjects had normal ability to identify and categorize the colors they perceived, they were asked to accept a pre-test to examine their knowledge about sorting colors to be at or above average level. Those subjects who did not participate in this study were color blindness or weak identification toward colors.

At training stage, we provided a conception formation test to help subjects set correct categorization for the colors recruited in the study. The target colors included: red, orange, yellow, brown, green, blue, purple, black, gray, and white. Subjects were presented a color square and a description on computer screen, and they needed to press "yes" or "no" button to confirm whether the description is correct. If they provided correct answer, screen would show a positive response, and a negative response was given when they provided incorrect key. The experiment operator evaluated the subject's performance. If subject's precision rate achieved to 95%, they could attend the main experiments. If they failed to get to the threshold, they had to take the concept exercise again, until they passed the training requirement.

3.1.2 Colors and Phonological Similarity

Phonological similarity was controlled as an independent variable in this study. The number of speech errors and the response time were dependent variables in this study. It was controlled to maintain colors of each trial in a state of phonological homogeneity. In the following experiments, we needed to sort the phonological relation for each color pair. Jaeger (2005) had evaluated the phonetic similarity between targets and error words for English children, and used phonological and phonetic criteria to judge whether the lexical errors could be considered phonologically related. It appeared to be an ideal basis for us to adapt and extend to evaluate the phonological similarity in Chinese system. Jaeger provided a list of criteria for grading the degree of similarity in English speech errors, which is provided in Table3-1. Jaeger supposed that a lexical error is considered "phonologically related" if its score achieved to at least 3 points. This evaluation system is also the criterion of phonological similarity in present study.

Table 3-1. Jaeger's Phonological Similarity Grading

Criteria for grading phonological similarity in lexical errors.
1. Same number of syllables = 1 point
2. Same stress pattern:
a) if both words are stressed, but each only one syllable = 0.5 points.
b) if both words have 2 or more syllables and the same number of syllables, and same primary stress location = 1 point.
c) if one word is one syllable and stressed, and the other is two-syllable and has first-syllable stress = 0.5 points.
d) if both words have 2+ syllables, but a different number of syllables, but the same general word-stress pattern = 1 point (e.g. 'dictionary' and 'library').
3. Same initial phoneme = 1.5 points.
4. Same final phoneme = 1 point.
5. Same primary-stressed vowel = 1.5 points (if only one of them occurs before /r/ or a nasal, only 1 point.)
6. Other same phonemes in same position in primary-stressed syllables = 1 point each.
7. Other same phonemes in same position in non-primary-stressed syllables = 0.5 points each.
8. Same phonemes in same position in primary-stressed syllable in one word, but non-primary stressed syllable in the other word, if syllables are in same position in word = 0.5 points each.
9. If a vowel is both word-initial and primary stressed, or both word-final and primary stressed = 2 points total.

According to Table 3-1, we noticed that most of these criteria reflected the structure in English phonology, such as multi-syllable pattern in lexicon and stress pattern in syllable, where as Chinese might not have them. Criterion 2 and 5 to 9 are stress-related grading for lexical errors, since stress assignment plays an important part in English syllable. However, in Chinese, we have pitch pattern within a syllable, that refers to the level tone and contour tones. English has multi-syllable lexicons where as Chinese is viewed as a mono-syllabic way. Therefore, criterion 2 and 5 to 9 might not be necessary for Chinese. Furthermore, tone is an important organizing structure in Chinese and thus there should be a system for it. The following is a revised list of evaluating criteria as shown in Table 3-2, the evaluation of which is hoping to reflect the sound patterns in Chinese.

Table 3-2. The Revised Criterions on Sound Similarity in Chinese

Strucrure	Shared Units	Score
A (syllable)	1. Whole Syllable (with tone) =Word	4.5-5.5
	2. Whole Syllable (without tone)	4-5
B (sub-syllable)	3. Syllable Structure	1
	4. Syllable Number	1
	5.Initial Onset	1-1.5
	6. Rhyme	0.5-1.5
	7. Bare Tone	0.5
C (phoneme)	8. Vowel	1
	9. Coda (G / N)	0.5
	10. Pre-nuclear Glide	0.5

In order to make the criteria to well explain the similarity between error and target in Chinese, Table 3-2 considered the feature of syllable structure and the numbers in Chinese, and it modified Jaeger's methodology when grading phonological similarity. For there exists typology diverse, the first difference is to simplify the way of syllable counting. We don't need to count syllable amount for Chinese. All the character in Chinese is phonologically monosyllabic, so we count the lexical pair sharing the same number in one point, especially for the monosyllabic case of color terms in Chinese. Second, tone is another issue that its phonemic significance should have independent criterion for grading. According to Jaeger's criteria, the monosyllabic words which share the same stress pattern, regardless of the vowel or syllable number factor, could get 0.5 points. Therefore, lexical pair which shares the same tone could get 0.5 points. Even though we know that syllable number and pitch pattern are the main differences between English and Chinese, Table 3-2 seems to help deal with the typological differences and grade for Chinese pairs.

With regard to the counting methodology for the similarity, Table 3-2 has another rationale from the one of Jaeger's (2005). There are three structures in Table 3-2: (A) syllable, (B) sub-syllable, and (C) phonemes. Each lexical pair should be graded according to the structural hierarchy (from A, B, to C), and each pair can only be graded within one structural criteria. Lexical pair could only be graded by the criteria in A, B, or C independently, and can't be graded in any two of the structures at the same time (except for syllable structure and syllable number). For example, if there is a pair of homophones, *xong* (means red or great); they could only gain the similarity score within the criteria (shared units) of structure A (syllable). The pair couldn't get any score from either structure B (sun-syllable) or C (phoneme). Sub-structure means that the lexical pair shares unit which is smaller than syllable and larger than segment. Criteria of sub-syllabic structure include syllable structure, syllable number, initial onset, rhyme, and bare tone. The remaining phonemic structure, such as vowel, coda, and prenuclear glide, should be attributed to structure C.

As to the specific phonological scores in Table 3-2, we mainly refer to Jaeger's (2005) phonemic and syllabic grading criteria. In structure C (phoneme), if the pair shares the same vowel, it could get 1 point in phonological similarity. Take the pair *huang2* [xwɑŋ35] (yellow) and *gao1* [kɑw55] (tall) as example in Table 3-3.

Table 3-3. Lexical Pair of *Huang2* and *Gao1* (yellow-tall)

Criterions	Factors	Score
4. Syllable Number	monosyllable	1 point
8. Vowel	[ɑ]-[ɑ]	1 point
Total	(Medium similarity)	2 points

Their monosyllable similarity could get 1 point according to the third criterion. In addition, there is not any similarity except for their vowel [ɑ], so it can also get 1 point. The total score of their phonological similarity is 2 points.

Regarding the score of coda, there are four phonemes which can serve at coda position in Chinese, such as [n],ŋ[], [j], and [w]. If the lexical pair shares the final coda, it can get 0.5 points in phonological similarity. The example can be seen in Table 3-4, the pair *hei* and *kwai*.

Table 3-4. Lexical Pair of *Hei1* and *Kwai4* (black-strange)

Criterions	Factors	Score
4. Syllable Number	monosyllable	1 point
9. Coda	[j]-[j]	0.5 points
Total	(Low similarity)	1.5 points

The pair *hei1* [xej55] (black) and *kwai4* [kwaj51] (strange) shares the same final consonant [j]. Beside 1 point the pair could get from the criterion of syllable number, another 0.5 points is gotten because of the shared final coda. The total of the lexical pair is 1.5 points. As the former mentioned, each pair should be graded within the same structural criteria. Criteria 8 and 9 should not overlap with criterion 6; that is, if the lexical pair doesn't get any score from the rhyme (criterion 6), then the pair is allowed to be graded under criterion either 8 or 9.

The issue on the status of prenuclear glide has been disputed in Chinese. The related argumentation can be referred to: prenuclear glide as onset's secondary articulation (Duanmu 1990, 2002), forming consonant cluster with initial (Bao 1990), as an element of rime (Jiang2001, Wang & Chang 2001), being in indeterminate ststus (Bao 1995, Huang 2001, Wan 1997), existing in independent structure (Shen 1992, 1993), supposing in approach of mora structure (Yip 2003, Ma 2003), and arguing under X-bar approach (van & Zhang, 2008). The status is not the main issue in present study. To avoid disputing, we follow Bao's framework (1990, 1995) to be a tentative assumption in this study. In this study, we tentatively assume that prenuclear glide might belong to the initial structure to form a consonant cluster in some Chinese syllables. However, on the surface representation, prenuclear glide sometimes interacts with vowel, such as /j-an/ into [jɛn], /w-əŋ/ into [woŋ], and /ɥ-an/ into [ɥɛn] etc. In the case of [j], [w], and [ɥ] interacting with vowel, we should go back to their underlying forms to give score to the lexical pair. Jaeger gave the same initial 1.5 points, which consists the part of the main first consonant and the following ones. In criterion 10, if the lexical pair shares the same prenuclear glide, the phonological score is valued 0.5 points. Therefore, the range of initial score could be 1 to 1.5 points. Basically, the pair could get at least 1 point for sharing the part of onset. If the pair also shares the following glide, it could get 1.5 points in total. The following is the details for the score of initial position.

Initial onset and rhyme belong to the case. In Chinese, the initial part includes a main consonant which initiates a syllable, and follows a prenuclear glide in some words. Based on Jaeger's measurement, if the pair shares the same initial which consists of single segment, it could get 1 point for their similarity. If the initial segment shares the same following glide, the pair could get 1.5 points as in Table 3-5 and 3-6.

Table 3-5. Lexical Pair of *Lü4* and *Lan2* (green-blue)

Criterions	Factors	Score
4. Syllable Number	monosyllable	1 point
5. Initial Onset	[l]-[l]	1 point
Total	(Medium similarity)	2 points

Table 3-6. Lexical Pair of *Hui1* and *Huo3* (gray-fire)

Criterions	Factors	Score
4. Syllable Number	monosyllable	1 point
5. Initial Onset	[xw]-[xw]	1.5 points
Total	(Medium similarity)	2.5 points

In Table 3-5, *lü4* [ly51] (green) and *lan2* [lan35] (blue) share the same initial [l] as well as syllable number, so they could get 2 points. In the case of two segments existing in initial, as in Table 3-6, *hui1* [xwej55] (gray) and *huo3* [xwɔ21] (fire) share the same segments [xw-] at initial, beside the 1 point they could get for syllable number, they could get the other 1.5 points for sharing onset as well.

Rhyme is another possible unit for a pair to share. Rhyme in Chinese includes central vowel and sometimes with a nasal or glide as coda. If the pair which is two open-syllable words and shares rhyme, the pair could get 1 point for owning the same vowel, as the pair of *zi3* (purple) and *zhi1* (weave) as examples in Table 3-7.

Table 3-7. Lexical Pair of *zi3* and *zhi1* (purple-weave)

Criterions	Factors	Score
4. Syllable Number	Monosyllable	1 point
6. Rhyme	[ɨ]-[ɨ]	1 point
Total	(Medium similarity)	2 points

In Table 3-7, *zi3* [tsɨ21] (purple) and *zhi1* [tʂɨ55] (weave) are the open-syllable words which share only the rhyme part, or attribute the unit to single vowel, so they could get 1 point for sharing the same apical vowel [ɨ]. If the close-syllable pair shares the rhyme, it could get not only 1 point for vowel part, but extra 0.5 points for sharing coda segments, as shown in Table 3-8.

Table 3-8. Lexical Pair of *bai2* and *hai3* (white-sea)

Criterions	Factors	Score
4. Syllable Number	monosyllable	1 point
6. Rhyme	[aj]-[aj]	1.5 point
Total	(Medium similarity)	2.5 points

Like the pair *bai2* [paj35] (white) and *hai3* [xaj21] (sea) in Table 3-8, it shares no parts but the rhyme segments [aj], so the total of this similarity is 1.5 points. Whether it is either the case of open syllable or close syllable, the pair should be graded under the B structure. That is to say, once the pair is attributed to share the rhyme unit, it couldn't be accumulated the score again under the criteria of C (phoneme structure). It means that the pair *bai2* and *hai3* can't get another score under criteria 8 and 9.

It's not easy to define the status of tone to belong to specific consonant or vowel in a syllable, as reviewed in previous chapter. One thing cannot be challenged is that tone is specified in lexical mode. It would be better to regard tone as sub-syllabic structure, because its suprasegmental characteristics. If lexical pair shares the same tone, it could get 0.5 points, as the pair in Table 3-9.

Table 3-9. Lexical Pair of *lan2* and *u2* (blue-nothing)

Criterions	Factors	Score
4. Syllable Number	monosyllable	1 point
7. Bare Tone	tone 35-tone 35	0.5 point
Total	(Low similarity)	1.5 points

As shown in Table 3-9, the tones of *lan2* [lan35] (blue) and *u2* [u35] (nothing) are the same, so the pair could get 1.5 points, which includes 1 point for their monosyllabic form and 0.5 points for their 35 tone.

The remaining syllable number and syllable structure are the ways to measure the similarity of the phonological and lexical structure, rather than of unit-related information in criteria 5 to 10. As to the syllable number, each character maps to single syllable. Color terms are the typical and special group, which is monosyllabic, in modern Chinese. All of the color pairs in this study can get basic 1 point under the criterion of syllable number, as well as the homophonous pairs of color.

There are five kinds of syllable structure in Chinese, such as V, CV, CVC, CCV, and CCVC. Among the structures, in present research, we focus on their underlying representations when we measure the phonological similarity for each lexical pair. Under this criterion, if the word pair shares the same syllable structure, it could gain one point at most, which accords to the criteria for single syllable in Jaeger's work. For example, *zi* [tsɨ21] (purple) and *lü* [ly51] (green) both belong to CV structure, so they could get 1 point for the same CV structure. In current study, if the structure involves postnuclear glide or nasal existing in coda position, CVC and CVN should be attributed to the same CVC structure. One of the reasons is that criteria 6 and 9 describe the segmental difference. The other reason is that postnuclear glide and nasal to the closed syllable in Chinese couldn't induce any phonological derivation, such as what prenuclear glide does during phonological derivation, between underlying and surface structures. Therefore, there is not any necessity to distinguish the ending difference when measuring similarity. One of the examples can be seen in Table 3-10.

Table 3-10. Lexical Pair of *huang2* and *huei1* (yellow-gray)

Criterions	Factors	Score
3. Syllable Structure	CCVC	1 point
4. Syllable Number	monosyllable	1 point
5. Initial Onset	[xw]-[xw]	1.5 points
Total	(High similarity)	3.5 points

In Table 3-10, *huang2* [xwɑŋ35] (yellow) and *huei1* [xwej55] (gray) are the case that shares not only initial consonants (1.5 points), single syllables (1 point), but also the CCVC structure (1 point). The total of the pair could get 3.5 points. As to the prenuclear glide which causes sound variation between underlying and surface structures, we would like to adopt the words' underlying phonological forms to measure the sound similarity, as former mentioned.

With regard to criteria within structure A (syllable), we have two kinds of conditions: one is the pair which shares the same syllable as well as tone, and the other which shares the syllable only. That is to say, if the lexical pair is two homophonous words, or the words that share the whole syllable except for tone, we should use structure A to measure the similarity for the pair. Take a pair of two homophones [ly51] as example, as shown in Table 3-11.

Table 3-11. Homophonous Pair of [ly51] (green-law)

Criterions	Factors	Score
1. Whole Syllable	Tone included	4.5 points
Total	(High similarity)	4.5 points

One of its lexical mappings can refer to color "green", and another can refer to the character meaning of "law." For their phonological similarity, they share everything but for their characters, so we attribute the homophonous pair to have 4.5 points. For details, 1 point is for initial and rhyme respectively, syllable number and structure (open syllable) could also get 1 point for each, and the remaining 0.5 points could be counted for the same tone.

The second condition is the pair shares the whole syllable except for tone, as in Table 3-12.

Table 3-12. Homophonous Pair of *lan2* and *lan3* (blue-lazy)

Criterions	Factors	Score
2. Whole Syllable	Without tone sharing	4.5 points
Total	(High similarity)	4.5 points

In Chinese, syllable /lan/ can refer to words which own the same syllable, such as [lan35] (blue) and [lan21] (lazy). The phonological similarity of the lexical pair could get 4.5 points in total, including 1 point for syllable number, another 1 point for the same onset [l], 1.5 points for its same rhyme [an], and the other 1 point for its sharing structure CVC.

Based on above criteria for measuring phonological similarity, we measured all the lexical pairs among these color terms in Chinese. There were 10 colors recruited in this study, and they were matched by pairs of two to grade and judge their degree of similarity. Table 3-13 is the similarity of colors which were paired to measure their phonological similarity. The following colors are abbreviated as: R (red), O (orange), Y (yellow), Br (brown), Ge (green), Bl (blue), P (purple), Bk (black), Ga (gray), W (white).

Table 3-13. Phonological Relations of Color Pairs

Similarity	Points	Color Pairs / Criteria					
High	5	R-Y 3 4 5 7 9					
	4	Bk-Ga 4 5 6 7	R-Br 3 4 6 10				
	3.5	R-Ga 3 4 5	Bl-W 3 4 7 8	Y-Ga 3 4 5			
	3	R-O 3 4 7 9	R-Bk 3 4 5	Ga-Br 347 10	Y-Br 349 10		
Medium	2.5	R-Bl 3 4 7	O-Bl 3 4 7	O-Br 3 4 9	R-W 3 4 7	O-W 3 4 7	Bk-Br 3 4 7
		W-Bk 3 4 9					
	2	O-Y 4 7 9	Y-Bk 4 5	Ge-Bl 4 5	Bl-Br 3 4	P-Br 4 5	Ge-P 3 4
Low	1.5	Y-Bl 4 7	Y-W 4 7	W-Gr 4 9	O-Bk 3 4	Bl-Bk 3 4	W-Br 3 4
	1	R-Ge 4	R-P 4	O-Ge 4	O-P 4	O-Ga 4	Y-Ge 4
		Ge-Bk 4	Ge-Ga 4	Ge-Br 4	Bl-P 4	Bl-Ga 4	P-W 4
		Ge-W 4	P-Ga 4	P-Bk 4	Y-P 4		

According to the score of similarity, we divided these color pairs into three groups, such as high, medium, and low groups in similarity. High group indicates the pair with high phonological similarity, while low group indicates the pair with low similarity. If lexical pairs could get more than 3 points, we attributed them into high group. If less than 2 points, pairs were attributed to low group. The remaining pairs, measured between 2 to 2.5 points, belong to the medium group. For example, the most similar phonological pair is *hong2* [xoŋ35] (red) and *huang2* [xwɑŋ35] (yellow), which can be seen in detail as following table.

Table 3-14. Lexical Pair of *hong2* and *huang2* (red-yellow)

Criterions	Factors	Score
3. Syllable Structure	CCVC	1 point
4. Syllable Number	monosyllable	1 point
5. Initial Onset	[xw]-[xw]	1.5 points
7. Bare Tone	Tone 35	0.5 points
9. Coda	[ŋ]-[ŋ]	0.5 points
Total	(High similarity)	4.5 points

As shown in Table 3-14, the color pair shares criteria 3 (the same underlying CCVC structure: 1 point), 4 (the same syllable number: 1 point), 5 (the same initial onset: 1.5 points), 7 (the same tone: 0.5 points), and 9 (the same coda: 0.5points). The total of this pair is 4.5 points, which exceeds the high level of 3 points, and the phonological relation is attributed to the high similarity group. Another pairs which are attributed to high similarity group are black-gray, red-brown, red-gray, blue-white, yellow-gray, red-orange, red-black, gray-brown, and yellow-brown. The detailed graded criteria of these pairs can be seen in Table 3-13. The total of 1 point could be the least similar pair in phonological relation, such as red-green, green-black, and blue-gray of the 16 pairs, as shown in Table 3-13. For example, the pair *zi3* [tsɨ21] (purple) and *hei1* [xej55] (black) is the case which shares no aspects but for their syllable number, so the pair could get 1 point.

In this study, we adopted the color pairs from high and low similarities into stimuli design because phonological similarity was controlled to examine the performance difference in experiments. Pairs from medium were excluded from the study. The pairs whose scores exceed 3 points or fell on 1 point were recruited in experiments. In total, there were 10 pairs from high similarity group and 16 from low groups to be arranged into color and naming tasks. The following are the details concerning how these color pairs were organized in tasks.

3.1.3 Test Items

Based on the lexical pairs graded by the criteria in Table 3-2, we arranged the color pairs by the groups of phonological similarity, and organized them into experiments. The following are series of tasks which were tested for different purposes: (1) Color Naming Test, (2) Color Reading Test, (3) Stroop Naming Test, (4) Homophonous Naming Test, and (5) Shared Unit Test. The above tests were designed with different visual tasks. The tasks 1, 2, and 3 adapted Stroop's experiments (1935) for a Chinese version. As reviewed in chapter 2, Stroop used color squares to have subjects name the color of carriers, and used color terms to have subjects not only read in one task, but also name the color of carrier terms. Stroop observed the response time that subjects reacted during different visual tasks. In present study, we asked subjects to read or name colors, and we would also like to induce them to produce speech errors by means of showing visual competition. In the following tasks, phonological similarity and the character frequency were carefully controlled when designing color pairs of each stimulus strip. Ten Chinese colors were the targets in these experiments, including *xong2* (red), *cheng2* (orange), *xuang2* (yellow), *lü4* (green), *lan2* (blue), *zi3* (purple), *hei1* (black), *bai2* (white), *zong1* (brown), and *hui1* (gray).

3.1.3.1 Task 1: Color Naming Test

In color naming test, subjects had to identify the colors of the carrier squares. Each trial consisted of eight color squares. In order to examine the effect of phonological similarity, the ten colors were arranged according to their phonological similarity in Table 3-13, and colors in one strip should be organized within the same similarity group. For example, (a) represents one of the trials from high similarity group in color naming square task.

(a) 

The eight colors in (a) were paired in the way of Table 3-13, and each of the pair should come from the high similarity group. The ones in (a) are black-gray, yellow-red, gray-red, and yellow- gray, and all the pairs are over 3 points. To maintain the sound similarity within each trial, all the colors of the trial were supposed to keep high phonological similarity, and should avoid making any two of the sounds among a trial form low similarity in pair. Such as *hei1* in the first pair, *hui1* in third, and *hong2* in second and yellow in forth, any pair of their similarity is at least 3 points, which leads to keep the whole trial in a phonologically similar status. In the color naming test, there were another nine trials organized in the same way of (a), and then each of the trials was reversed to have another ten trials in high similarity relation. In total, there were 20 trials of high phonological similarity for subjects to name the colors of colored squares in each strip. The following (b) is one of the other trials with less phonological similarity.

(b) 

In trial (b), there are also four pairs of color squares which are from low similarity group, such as the pairs of green-white, yellow-purple, green-yellow, and orange-green. The scores of these pairs are 1 point in similarity, and any two colors of the cross-pairs should not exceed 2 points ideally. Even though the colors white and orange in (b) do not belong to the pair of low similarity, they do not reach at high-similarity level, either. With the goal to maintain the low similarity in sounding within the whole trial, any two of the color carriers should avoid the sound relation forming high similarity pair. There were also ten strips designed as trials of low similarity, and they were also reversed to act as the other ten trials in the naming test. There were twenty trials in total of low similarity.

Therefore, there were 40 trials in this naming task. Twenty of them were the ones of high similarity, and the other twenty trials were of low similarity. The actual trials recruited in the color naming test represent in appendix 1.

As to the procedure, subjects first saw a star mark in the center of the screen for 2 seconds, which hints the coming stimulus. When the colored trial showed up, from left to right, subjects had to name the colors for each square serially. The answer span was set 10,000 ms, and subjects needed to name all the colors without correcting backwards within this time limit. After naming all the squares, subjects needed to press the response button to have their response time record in computer as well as to activate the next trial. During the naming test, all the sound responses which subjects provided were collected with a SONY-IC recorder for later transcription and error analysis. We used E-Prime 2.0 as experiment software to design the whole experiments through this study, and to record the response time that subjects spent on each trial. After subjects finished a trial, in this study, they had to press the response key (SPACE button on keyboard) immediately, so that could help record the answering span precisely. In order to make subjects familiar with the whole procedure and get ready with the experimental equipments, we designed a practice session for them to get practice until they could complete each trial on their own.

3.1.3.2 Task 2: Color Reading Test

In color reading test, color terms were presented as visual stimuli at this phase. We asked subjects to read the character in the screen. All the characters were shown in color black, and as in the naming test, each trial was provided eight color terms with the manner of pairing. For the sake of counterbalanced, the color terms in reading test were substituted for the corresponding color carriers in naming test. The four lexical pairs should be attributed to the same similarity group, as in (c) and (d).

(c) 黑灰 黃紅 灰紅 黃灰

'Hei1-hui1 huang2-hong2 hui1-hong2 huang2-hui1'

'Black-gray yellow-red gray-red yellow-gray'

(d) 綠白 黃紫 綠黃 橙綠

'*Lü4-bai2 huang2-zi3 lü4-huang2 cheng2-lü4*'

'Green-white yellow-purple green-yellow orange-green'

Trial (c), which the colors of squares in (a) correspond to, was set to test the pairs from high similarity group, and subjects needed to read the character they saw, not to name the color of it. All the color pairs were in homogeneity of high similarity, and any two of terms in a trial should avoid being in low similarity relation, designed as the same way in color naming test. Trial (d) is the case of low similarity in phonological relation. All the pairs were paired in less similarity, and were avoided any two colors of the trial from forming phonological relation of high similarity.

All the trials were designed in the same way as those in the color naming test, but shown in color terms. Subjects would first saw a star marker in the center of the screen for two seconds, four color pairs followed up, and had endured for 20,000 ms for subjects to read the trial. There were twenty trials which were designed for high phonological group, and the other twenty were in low similarity relation. All the trials in this session are displayed in appendix 1.

During this reading test, both the sound sample and the response temporal data were record by SONY-IC recorder and computer respectively. E-prime helped to collect and record the span between the start of the trial and the SPACE pressing which put answering to an end. Practice session was also recruited to have subjects get ready to complete the trials before experiment.

3.1.3.3 Task 3: Stroop Naming Test

A character can carry two semantic ideas about color at the same time: the messages from the color term (literal information) and its ink color (visual information). Most of the time, we don't need to name the colors of the term when reading word. Stroop's technique provided a phonological competition in mind for subjects to ignore the effect of reading intentionally, and to name the colors of the serial colored carriers instead. In this study, we adopted Stroop's experiment, and we substituted the Chinese color terms for the original ones. In order to know better whether phonological similarity is facilitation or inhibition in lexical processing, we still divided the color pairs into high, medium, and low similarity groups, as the former naming and reading test.

In the Stroop naming test, color terms are again to be the visual stimuli, including the ten colors as recruited in former tests. Howerer, subjects were asked not to read the color terms in a trial, but to name the colors of the characters serially. The color pairs as well as their order in each trial of previous naming and reading tests were kept in this session, but the ink color of each character was incompatible with the term referred to. Both the colors in a pair were exchanged their ink color, so each color of a term pair was the other term referred to. The visual terms, whose color is always inconsistent, were displayed in the same manner as in previous tests, but the sequence of the trials was random in order to avoid practice effect from the first naming task to this session. Take (e) and (f) as samples, whose carriers are the same in (c) and (d):

(e) 黑灰 黃紅 灰紅 黃灰

'Hei1-hui1 huang2-hong2 hui1-hong2 huang2-hui1'

'Black-gray yellow-red gray-red yellow-gray'

(f) 綠白 黃紫 綠黃 橙綠

'Lü4-bai2 huang2-zi3 lü4-huang2 cheng2-lü4'

'Green-white yellow-purple green-yellow orange-green'

The four lexical pairs in (e) are the ones with high phonological similarity on literal representation. However, the ink colors of each pair were exchanged. In each pair, the color of first carrier is indicated by the other term. For example, the color of the first term in "black-gray" is assigned to be gray, and the second one is assigned black by the first term. The remaining three pairs of a trial were organized in the same manner. In (f), the lexical pairs were in low phonological relation. The way of the term and their color representation was the same with (e). The pair "yellow-purple" was painted in purple for the first and in yellow for the other term. The total amount of the trials in this session is the same with the previous tests. There were twenty trials with phonological relation, and the other twenty were in less relation of phonological similarity. All the trials in this session are displayed in appendix 1.

Before entering the experiment, subjects have to take exercise phase, and they should not only get familiar with the procedure but also know that visual color was the most concerned in this section, rather than the terms. All of the procedure was repeated as the former tests, subject would first gaze at the star mark for 2000 ms, and the trial showed up in the fallowing 20 seconds. They had to name the color for each of the pair within the time limit. They were asked to press the response key right after they finished answering, or the star hint of the next trial would show up automatically to inform subjects the coming trial. Their naming response and their reacting time were also recorded in the same way at this phase.

3.1.3.4 Task 4: Homophonous Naming Task

From above visual tasks, we would like to know how the sound similarity and visual representation induced effect on lexical process, which would reflect on the speech errors and the response time. In this section, we recruited a set of homophones as a substitution for color terms. Homophonous Stroop naming test was designed to take homophones instead of color terms to observe whether the lexical process would be similar to that for color terms. With regard to the test material, the target colors were still the same with previous experiments. The colors included red, orange, brown, yellow, green, blue, purple, gray, white, and black again, but the visual stimuli were represented as the corresponding homophones of those colors. The organization and order of stimuli was the same with Stroop naming test. In order to make sure that every aspect kept similar, except for the semantic domain, when we used homophonous characters as visual stimuli, we had to control the frequency of the characters to have them persuasible to replace for the color terms. Even though it is impossible to find the homophones sharing the same frequency with their corresponding colors, we could find a set of homophones whose distribution and correlation of word frequency was similar to those among color terms. The following Table 3-15 is the homophones of color terms and their character frequency in Chinese.

Table 3-15. Word Frequency of Color Terms and Corresponding Homophones

Color Term	紅	棕	橙	黃	綠	藍	紫	灰	白	黑
Meaning	red	brown	orange	yellow	green	blue	purple	gray	white	black
Frequency[3]	503	11	11	367	200	168	30	116	661	437
Homophone	洪	鬃	懲	皇	慮	蘭	秄	輝	白	嘿
IPA	xoŋ35	tsoŋ55	tʂʰəŋ35	xwɑŋ35	ly51	lan35	tsɨ21	xwej55	white	xej55
Frequency	23	3	1	**	10	10	6	3	661	106

** indicates there is no data or information

Table 3-15 shows each word frequency for each color term in Chinese. The number indicates the specific times which the word appears in language. The most frequent is white (N=661), and follows red (N=503), black (N=437), yellow (N=367), green (N=200), blue (N=168), gray (N=116), purple (N=30), brown (N=11), and orange (N=11). The words in lower row are the corresponding homophones and their frequency. These are some missing statistics in the homophone candidates. For example, [xwɑŋ35] is one of the case that there is no other characters, except for "黃", having frequency statistics in the Sinica Corpus. With an ad hot approach, we choose "皇 (experor)" as a corresponding homophone, because it is the most transparent literally in reading task among all the candidates. Another missing data is the homophone of color white "白 [paj35]." In Chinese, there is not any character which owns the same tonal syllable. To maintain the whole trial in a homophonous condition, we still recruited the term white again in this homophonous naming task.

However, it seems impossible to find a set of homophonous words with absolutely the same frequency. With regard to keep frequency balanced both among the color terms and their homophones, we needed to have these frequency statistics from Sinica Corpus (2005) to pass the examining of correlations test. We had to

[3] The frequency data were quoted from Cheng, Huang, Lo & Tsai (2005) "Word List with Accumulated Word Frequency in Sinica Corpus."

examine the frequency relation among the homophones we had used to compare with the relation among the color terms, so that we could make sure that frequency bias would induce little influence in naming test. Table 3-16 shows the correlation test between color terms and their homophonous counterparts.

Table 3-16. The Frequency Correlation between Color and Homophone

Descriptive Statistics	Mean	Std. Deviation	N
Color term	175.75	178.65	8
Homophone	8.00	7.48	7

Correlations		Color term	Homophone
Color term	Pearson Correlation	1	.961**
	Sig. (2-tailed)		.001
Homophone	Pearson Correlation	.96**	1
	Sig. (2-tailed)	.001	

**. Correlation is significant at the 0.01 level (2-tailed).

From Table 3-16, the mean scores of the two variables are 175.75 and 8.0 respectively, while the ones of standard deviation are 178.65 and 7.48 respectively. The Pearson correlation is .96 ($p<.05$), which reaches to the level of significance. That is to say, the frequency distribution among the color terms correlates the distribution among the homophones positively and significantly ($r=.961$, $p<.05$).

The correlation examination shows that the homophone carriers keeps the character frequency balanced with color terms. The trial designs in Stroop's naming test were repeated in homophonous naming task, except for the homophones in Table 3-15 which were substituted for the original color stimuli, as shown in sampler trials (g) and (h).

(g) 黑藍 橙紫 綠黑 黃藍

(h) 嘿蘭 懲籽 慮嘿 皇蘭

The sampler (g) was used in Stroop's naming test, and (h) is the corresponding trial in the homophonous Stroop's naming task. All the lexical pairs in (h) were replaced by the homophones in Table 3-15, and the visual colors were arranged in the same order with (g). Of the trial in (g), the color pair "黑藍 [xej55 lan35]" (black-blue) was replaced with homophonous pair "嘿蘭." Their visual colors were painted in blue for the former and black for the other respectively. The other pairs orange-purple, green-black, and yellow-blue were replaced in the same manner, as shown in (h). The other phonological trials of this experiment are displayed in appendix 1.

During this phase, there was nothing different from Stroop's task to name the color of lexical, except for the homophone stimuli. After the gazing at star mark for two seconds, subjects were asked to name the phonological pairs for each trial, and then pressed the response key to have their answers as well as reacting span be recorded. The total response time was limited within 20000 ms, and the next trial would show up automatically.

3.1.4 General Procedure and Data Analysis for Task 1 ~ 4

At the beginning of the testing phase, subjects could saw a star mark in the center of the screen for 2000 ms, which reminded subjects to get ready for the coming trial. Then the visual stimulus was displayed, and subjects were asked to name the visual color for each of the carriers. During the answering period, the SONY-IC recorder was set aside to collect the answers which subjects provided. After finishing answering, they had to push the response button on the serial response box to record the reaction time as well as to initiate the next trial. When answering a trial, subjects were asked to gaze at a single carrier and process them serially and one by one. It was prohibited to change, skip, reverse, or omit the processing order of carriers in a single trial. The above is a general course for completing a trial. As to the amount of trials, in experiment 1 and 2, forty trials were designed in color square naming test. Twenty of them were composed of the ones with high phonological similarity, and the others were of the ones with low phonological similarity. In the following term reading test and Stroop's naming test, the same trials in the former test were applied again to different goals of individual tasks. Therefore, there were also 40 trials in term naming test, and so are in color reading test, Stroop's naming test and homophonous naming test. In total, 160 trials were displayed to 22 subjects, 3,520 trials could be collected in experiment 1 and 2. That is to say, there were 28,160 tokens could be thought to be generated (8 carriers in each trial). All of the errors will be graded and judged in terms of phonological similarity, and the phonological relation between target and error will be analyzed and counted for later discussion. Three types of the relationship between target and error were concerned in the two experiments, such as phonological errors, semantic errors, and mixed errors. Regarding to the phonological and mixed errors (both are phonologically related errors), six types of phonologically structural relations were analyzed for them, such as the structures of onset, vowel, rhyme, syllable, tone, and tonal syllable.

E-prime (a program for experiment design and operation) helped us to record the response time (RT) that subjects had done for trials. The response span among the four tests will be compared, and the temporal patterns would be compared to the pattern of speech errors in the following discussion.

3.2. Task 5: Shared Units Naming Task

By means of evidence in speech error, we want to know whether phonological units could imply significant effect, and what structural representations would play important role during word encoding. In this section, we also performed a naming experiment, but the visual stimuli were not color terms any longer. We designed some phonological groups of word lists, which shared certain kind of phonological structures with the color, to substitute for color terms.

In order to find a set of words to replace for the terms and to avoid the word frequency effect meanwhile, we had to choose the corresponding characters according to their frequency distribution from the "Word List with Accumulated Word Frequency in Sinica Corpus." The shared unit task was designed to examine the issue of planning unit in word production. According to the phonological structure in Chinese syllable, we had six kinds of units to organize for the shared unit test in this study, which included onset, vowel, rhyme, bare syllable, bare tone, and tonal syllable (syllable + tone). The following will divided into six parts to describe the word designs of these structural representations. In order to look into how the segments are encoded into a phonological structure and how they interact within sound competition in visual perception, we excluded the factor of syllable structure in present study, though it is still important on the issue of the syllable unit in speech production to be chunk or schema.

3.2.1. Shared Unit: Onset

As reviewed in previous literature chapter, onset seems to be more likely to slip than non-initial ones (Dell & Juliano & Govindjee 1993). In this section, we would like to design a set of words which only share the part of onset with the color terms to substitute for the original visual terms. Table 3-17 lists the corresponding characters as well as their word frequencies, and their correlation analysis is given in Table 3-18.

Table 3-17. Word list of onset sharing

Color Term	紅	棕	橙	黃	綠	藍	紫	灰	白	黑
IPA	xoŋ35	tsoŋ55	tʂʰəŋ35	xwɑŋ35	ly51	lan35	tsɨ21	xwej55	paj21	xej55
Meaning	red	brown	orange	yellow	green	blue	purple	gray	white	black
Frequency	503	11	11	367	200	168	30	116	661	437
Onset Sharing	忽	宰	侈	毀	爐	撈	卒	禍	補	壺
IPA	xu55	tsaj21	tʂʰɨ21	xwej21	lu35	lɑw55	tsu35	xwɔ51	pu21	xu35
Meaning	sudden	control	luxury	destroy	stove	scoop	soldier	trouble	cram	pot
Frequency	76	29	1	65	38	23	8	45	109	65

In Table 3-17, the lower row is the words that share the onset part with the corresponding color terms in upper row, as well as their respective frequency data. Take the color red [xoŋ35] as example, "忽 [xu55]" is a word which shares only the onset segment [x]. Other aspects such as vowel, coda, tone and its phonological structure don't overlap with those in the target term. Another example is the color green [ly51], in the case of the open syllable. We took the word "爐" [lu35] as the corresponding one because it shares the onset segment, without overlapping vowel and its tone. As to the compatibility of word frequency among the chosen characters, we examine their correlation with those of color terms by means of correlation test.

The mean scores of the two variables are 2.50 and 46.0 respectively, while the scores of standard deviation are 228.86 and 33.09 respectively. The Pearson correlation is .948 ($p<.05$), which reaches to the level of significance. It appears that the frequency distribution among the color terms correlates the distribution among the onset carriers positively and significantly ($r=.948$, $p<.05$).

Therefore, we adopted the words in Table 3-17 to substitute for the color terms of the naming test. For the actual stimuli design, we used the same sequence of the color trials in the former experiments into this phase, but two different rationales were applied in the design of this section. First, we don't need to pair those color carriers as disyllable words. The eight characters were represented independently within a trial. Second, after arranging the characters, we painted them in the target color of the carrier, instead of painting the term in another color in previous tests. For example, the following trial (i) is one of the samples in this unit sharing test and we painted the carriers in their target colors, while (c') is the trial from the color reading test and characters were painted as the term indicated. We arranged the target colors in the same way as in (c') for the colors of trial (i).

(c') 黑灰黃紅灰紅黃灰

[xej55 xwej55 xwɑŋ35 xoŋ35 xwej55 xoŋ35 xwɑŋ35 xwej55]

'Black gray yellow red gray red yellow gray'

(i) 壺禍毀忽禍忽毀禍

[xu35 xwɔ51 xwej21 xu55 xwɔ51 xu55 xwej21 xwɔ51]

'Pot trouble destroy sudden trouble sudden destroy trouble'

Trial (c') and (i) were painted in the same colors for each of the positions. We didn't pair them into four phonological words as in (c). The visual color sequence of (i) was black, gray, yellow, red, gray, red, yellow, and gray. The color sequence not only accorded to the color terms in (c'), but also applied to the shared unit carriers of (i)—pot (壺), trouble (禍), destroy (毀), sudden (忽), trouble (禍), sudden (忽), destroy (毀), and trouble (禍).

Subjects still had to name the colors of those carriers, instead of reading the carrier words. There were 20 trials in total in this onset sharing section. All of the trials (20 in total) came from the former naming and reading experiments, but we didn't reverse them to have another twenty. Among them, ten trials were recruited from the phonological group of high similarity, while the other ten came from the group of low similarity. In addition, the amount and target colors of trials in the following vowel sharing, rhyme sharing, syllable sharing, tone sharing, and tonal syllable sharing tests were organized in the same way. The differences among these unit sharing tests were not only the sharing structures, but also the individual set of word carriers for each unit.

In this shared unit naming task, we wanted to compare the naming precision and the speed of naming among the phonological units. In this part, we would like to know whether the shared unit, onset part, induced any effect on lexical processing and the pattern of speech errors as well as the amount.

3.2.2. Shared Unit: Vowel

Vowel is the center part in a syllable, and syllable in Chinese doesn't exist without vowel segment. Vowel seems to be the essential material for syllable constituting. Table 3-18 is the set of words which were used to replace for the color terms. The substituting words and color terms contrast for single vowel.

Table 3-18. Word list of vowel sharing

Color Term	紅	棕	橙	黃	綠	藍	紫	灰	白	黑
IPA	xoŋ35	tsoŋ55	tʂʰəŋ35	xwɑŋ35	ly51	lan35	tsɨ21	xwej55	paj21	xej55
Meaning	red	brown	orange	yellow	green	blue	purple	gray	white	black
Frequency	503	11	11	367	200	168	30	116	661	437
Onset Sharing	狗	剖	奮	告	徐	擦	釋	**	反	**
IPA	kow21	pʰow21	fən51	kɑw51	ɕy35	tsʰa55	ʂɨ51	**	fan21	**
Meaning	dog	dissect	excite	sue	slow	wipe	explain	**	contrary	**
Frequency	380	3	2	116	94	75	18	**	629	**

** indicates no data.

Take color blue as a sample, the vowel of its surface form is [a], and we found a corresponding carrier which was vowel sharing, "擦 [ts^{h}a55]" (wipe). Vowel sharing should exclude the involvement of rhyme. For example, the rhyme of color brown is [oŋ], and we avoided a corresponding carrier which shared vowel as well as coda, except for open syllable of green and purple. The character "剖 [p^{h}ow21]" (dissect) was a rather proper candidate to substitute for brown, because it only shared the vowel [o], rather than the whole rhyme. In this set of wordlist, we noticed there are two gaps because there are no proper correspondences to replace for the color gray and black in Chinese. To keep the whole shared unit naming task balanced in trial designing and effective for later analysis, we still kept the gap filled with original gray and black to make the target colors of trials in this section the same with the ones in the other shared unit sections. Therefore, in this vowel test, we still used "灰" (gray) and "黑" (black) in trials. Through correlation test of word frequency between color terms and vowel sharing carriers, the mean scores of the two variables are 2.50 and 1.64 respectively, while the scores of standard deviation are 228.86 and 224.08 respectively. The Pearson correlation is .940 ($p<.05$), which reaches to the level of significance. It shows that the frequency distribution among the color terms correlates the distribution among the vowel carriers positively and significantly ($r=.940$, $p<.05$). Through the examination of significance, we adopted the word list in Table 3-18 to be the vowel carriers for their corresponding colors. The sequence of target colors was the same as in the former color reading test. Besides, we didn't pair the carriers within trials, as represented in onset test. The following (d') and (j) are the sample trials in this section.

(d’) 綠 白 黃 紫 綠 黃 橙 綠

[ly51 paj35 xwɑŋ35 tsɨ21 ly51 xwɑŋ35 tʂʰəŋ35 ly51]

‘Green white yellow purple green yellow orange green’

(j) 徐 反 告 釋 徐 告 奮 徐

[ɕy35 fan21 kɑw51 ʂɨ51 ɕy35 kɑw51 fən51 ɕy35]

‘Slow contrary sue explain slow sue excite slow’

The vowel carriers of (j) were one of the trials which were used to compare the performance of the colors in (d’). As to such case, the color sequence in both trials was green, black, yellow, purple, green, yellow, orange, and green. This visual color sequence was applied to the vowel carriers in (j), such as 徐 *xu2* (in green), 反 *fan3* (in black), 告 *gao4* (in yellow), 釋 *shi4* (in purple), 徐 *xu2* (in green), 告 *gao4* (in yellow), 奮 *fen4* (in orange), 徐 *xu2* (in green). The other trials of vowel can be seen in appendix 2. In this part, we would like to know whether the shared unit, vowel part, induced any effect on lexical processing and the distribution of speech errors.

3.2.3. Shared Unit: Rhyme

Rhyme composes of nucleus and coda in a syllable. In Chinese, both V and VC structure are patterns of rhyme, which constitute open and closed syllable respectively. For some trials in shared unit test, we recruited a series of character carriers which shared the part of rhyme to be the substitution of color terms in naming task. Table 3-19 displays the wordlist of rhyme sharing characters.

Table 3-19. Wordlist of rhyme sharing

Color Term	紅	棕	橙	黃	綠	藍	紫	灰	白	黑
IPA	xoŋ35	tsoŋ55	tʂʰəŋ35	xwɑŋ35	ly51	lan35	tsɨ21	xwej55	paj21	xej55
Meaning	red	brown	orange	yellow	green	blue	purple	gray	white	black
Frequency	503	11	11	367	200	168	30	116	661	437
Onset Sharing	冬	拱	亨	趟	距	閃	癡	備	海	累
IPA	toŋ55	koŋ21	xəŋ 55	tʰɑŋ51	tɕy51	ʂan21	ʂɨ55	pej51	xaj21	lej51
Meaning	winter	arch	henry	trip	distance	flash	idiotic	prepare	sea	tired
Frequency	342	3	3	167	119	81	14	70	613	233

For example, “閃 [ʂan21]” (flash) is one of the candidates which shares the part of rhyme [an] with the term *lan2* (blue). Another case is the term with open syllable. In vowel sharing section, color green and purple are the cases of open syllable, but they also attribute to the case of rhyme structure. Therefore, for rhyme sharing, we still had to recruit other carriers to replace for open syllable colors, such as “距 [tɕy51]” (distance) and “癡 [ʂɨ55]” (idiotic) in Table 3-19.

As to the frequency distribution among carriers, we examined the frequency correlation between color terms and rhyme sharing carriers.

The mean scores of the two variables are 2.50 and 1.65 respectively, while the scores of standard deviation are 228.86 and 191.53 respectively. The Pearson correlation is .949 ($p<.05$), which reaches to the level of significance. It shows that the frequency distribution among the color terms correlates the distribution among the rhyme sharing carriers positively and significantly ($r=.949$, $p<.05$).

With the significance in the bi-tailed correlation test, we recruited the carriers in Table 3-19 to serve as color substitution. The basic trials were kept the same with the ones in former sharing units, except for the replacing carriers.

(c’) 黑灰黃紅灰紅黃灰

[xej55 xwej55 xwɑŋ35 xoŋ35 xwej55 xoŋ35 xwɑŋ35 xwej55]

‘Black gray yellow red gray red yellow gray’

(k) 累備趟冬備冬趟備

[lej51 pej51 t^hɑŋ51 toŋ55 pej51 toŋ55 t^hɑŋ51 pej51]

‘Tired prepare trip winter prepare winter trip prepare’

Trial (k) is the corresponding one for trial (c’) in this rhyme task. All of the carriers in trial (k) shared the rhyme part with the colors in (c’). For instance, color gray in Chinese is pronounced as [xwej55], and its corresponding carrier “備” (prepare) is [pej51]. They shared the part of rhyme. The target colors that were applied to (k) were the ones in (c’). Subjects needed to name the ink colors painted for the carriers in (k), and, that is, the color terms in (c’) was the answers to trial (k). The trials of color terms, such as (c’), were sometimes intervened into the shared unit naming phase, but the main concern of this phase is the performance in speech error and reacting span for the unit sharing trials, such as (k).

3.2.4. Shared Unit: Bare Syllable

When the tone is excluded from the sharing part, we can say that the word pair shares the bare syllable. In this naming section, we recruited a set of characters which share the part of bare syllable to be the carriers for colors. Table 3-20 displays the set of carriers as well as their word frequency.

Table 3-20. Wordlist of sharing bare syllable

Color Term	紅	棕	橙	黃	綠	藍	紫	灰	白	黑
IPA	xoŋ35	tsoŋ55	tʂʰəŋ35	xwɑŋ35	ly51	lan35	tsɨ21	xwej55	paj21	xej55
Meaning	red	brown	orange	yellow	green	blue	purple	gray	white	black
Frequency	503	11	11	367	200	168	30	116	661	437
Onset Sharing	哄	粽	逞	慌	驢	覽	滋	迴	敗	**
IPA	xoŋ21	tsoŋ51	tʂʰəŋ21	xwɑŋ55	ly35	lan21	tsɨ55	xwej35	paj51	**
Meaning	uproar	rice dumpling	show off	flurry	donkey	view	nourish	circle	lose	**
Frequency	34	1	1	26	24	17	1	3	123	**

The words in Table 3-20 are the candidates that share the bare syllable with colors. For instance, the word "迴 [xwej35]" (circle) shares the part of the whole syllable except for tone with color gray [xwej55]. Therefore, we call that the color and the carrier share the unit of bare syllable. However, there is a lexical gap for the counterpart of color black "黑 [xej55]" in Chinese. We can not find any word for color black that only shares bare syllable without tone. Since the missing slot existed, we still had to examine the frequency correlation by means of bi-tailed correlation test.

The bi-tailed correlation between color terms and their corresponding carriers achieved the level of significance. The mean scores of the two variables are 2.50 and 2.56 respectively, while the scores of standard deviation are 2.29 and 3.87 respectively. The Pearson correlation is .878 ($p<.05$), which slightly reaches to the level of significance, although the set of words were the most closed to the value of .001 level with post hoc checking from database. It shows that the frequency distribution among the color terms still correlates the distribution among the bare syllable carriers positively and significantly ($r=.878$, $p<.05$).

It seems that the carriers in 3-20 might be the proper ones for color replacement so far. In order to strike a balance with former tasks, except for the carriers, the color arrangement and the target colors within each trial were kept the same with previous tasks. Trial (l), varied from (d’), an example of this part.

(d’) 綠白黃紫綠黃橙綠

[ly51 paj35 xwɑŋ35 tsɨ21 ly51 xwɑŋ35 $tʂ^{h}$əŋ35 ly51]

‘Green white yellow purple green yellow orange green’

(l) 驢敗慌滋驢慌逞驢

[ly51 paj35 xwɑŋ35 tsɨ21 ly51 xwɑŋ35 $tʂ^{h}$əŋ35 ly51]

‘Donkey lose flurry nourish donkey flurry show off donkey’

The carriers in trials (d’) and (l) shared the phonological part of the whole bare syllable. In this wordlist, we would like to recruit a set of words which had not any semantic relation but shared the same syllable in this experiment. For example, color green in Chinese shares the same bare syllable with the word “驢 [ly35]” (donkey). They both shared the part of syllable [ly]. Subjects were asked to name the ink colors painted for the carriers in (l), and, that is, the color terms in (d’) was the answers to trial (l). The trials of color terms, regular mapping such as in (d’), were sometimes intervened into the ones during this shared unit naming phase, but the main concern of this phase is the performance in speech error and reacting span for the unit sharing trials, such as the colors and carriers in (l).

3.2.5. Shared Unit: Bare Tone

Tone is the use of pitch in language to distinguish lexical or grammatical meaning, although tone plays little role in modern Chinese grammar. In Chinese, tone includes four lexical tones (one level tone and three contour tones) and a nurture tone.

For example, "屋 [wu55]" (house), "無 [wu35] " (nothing), "五 [wu21]" (five), and "霧 [wu51]" (fog) are the ones which share everything in syllable except for their tones, and those tones lead to meaningful contrast. Even though tone is not a phoneme, it acts phonemically. In this tone sharing section, we hired a set of words which share nothing in syllable but the lexical tone. Table 3-21 is the wordlist of the carriers which share lexical tone with colors as well as their frequencies.

Table 3-21. Wordlist of sharing tone

Color Term	紅	棕	橙	黃	綠	藍	紫	灰	白	黑
IPA	xoŋ35	tsoŋ55	tʂʰəŋ35	xwɑŋ35	ly51	lan35	tsɨ21	xwej55	paj21	xej55
Meaning	red	brown	orange	yellow	green	blue	purple	gray	white	black
Frequency	503	11	11	367	200	168	30	116	661	437
Onset Sharing	離	苛	鄒	凡	敗	足	普	貪	初	屋
IPA	li35	kʰɤ	tsow55	fan35	paj51	tsu35	pʰu21	tʰan55	tʂʰu55	u55
Meaning	leave	harsh	surname	any	lose	foot	universal	greedy	frst	house
Frequency	349	2	1	220	123	102	13	60	616	270

The carriers in Table 3-21 are the candidates which were substituted for the color terms. Each of them shares only lexical tone with their corresponding color. Take the carrier "屋 [wu55]" (house) as an example, the target color is black "黑 [xej55]." In terms of the syllable structure and segment, they only share the level tone 55. To strike a balance between the syllable structure and the following examination of word frequency, "紫 [tsɨ21]" (purple) and "普 [pʰu21]" (unverial) share not only the tone 21, but also the phonological structure CV. Even though phonological syllable could be a possible unit for planning, we still chose to ignore and focused on the control of the word frequency. The other candidates were not the cases of sharing syllable structure. Even though "橙 [tʂʰəŋ]" (orange) and the intended substitution "鄒 [tsow55]" (surname in Chinese) are both attributed to general CVC structure, CVC and CVN were not precisely the same in terms of syllable pattern.

The bi-tailed correlation between color terms and their corresponding carriers achieved the level of significance. The mean scores of the two variables are 2.50 and 1.76 respectively, while the scores of standard deviation are 228.86 and 195.07 respectively. The Pearson correlation is .971 ($p<0.05$), which reaches to the level of significance. It shows that the frequency distribution among the color terms still correlates the distribution among the lexical tone carriers positively and significantly ($r=.971$, $p<.05$).

With the significance in the bi-tailed correlation test, we could make sure the carriers in Table 3-21 could be the substitution for color terms in this section. The carrier trials and the sequence of the carriers were kept the same with the previous sections, except for the replacing carriers. Trial (m) is one of the trials displayed below as a sample.

(c') 黑灰黃紅灰紅黃灰

[xej55 xwej55 xwɑŋ35 xoŋ35 xwej55 xoŋ35 xwɑŋ35 xwej55]

'Black gray yellow red gray red yellow gray'

(m) 屋貪凡離貪離凡貪

[wu55 t^han55 fan35 li35 t^han55 li35 fan35 t^han55]

'House greedy any leave greedy leave any greedy'

In this section, each trial was arranged in a sequence of eight carriers. The visual words in trial (m) share the lexical tone with the corresponding colors in (c'). For example, "屋 [wu55]" (house) shares the tone with "黑 [xej55]" (black), "貪 [t^h an55]" (greedy) shares the level tone with "灰 [xwej55]" (gray), "凡 [fan35]"

(universal) shares the tone 35with “黃 [xwɑŋ35]” (yellow), and “離 [li35]” (leave) shares the tone 35 with “紅 [xwoŋ35]” (red). Subjects still had to name the colors of those carriers, instead of reading the carrier words. There were 20 trials in total in this tone sharing section. Among them, ten trials were again recruited from the phonological group of high similarity, while the other ten came from the group of low similarity. We would like to mainly concern whether tone is really separable and independent from syllable in phonological representation or lexical mode, as mentioned in the work of J. Y. Chen (2002).

3.2.6. Shared Unit: Tonal Syllable

When two words share the whole syllable (tone included), except for their meaning, we call the word pair is in homophonous relation. In this study, orthography was also needed to be taken into consideration. We used a set of characters which shared the whole syllable with colors, and asked subjects, again, to name the colors of the carriers, rather than reading the term. It seems similar that we did in the former Homophonous Stroop’s Task, but we varied some of the ways in this shared unit section. First, all the carriers were not paired in a trial. They were arranged in a sequential way. We asked subjects to name the colors as reading a sentence without pause. Second, most of the carriers were painted in their target colors, such as the carrier “慮 [ly51]” (consider) was painted in green ([ly51] in Chinese). However, subjects were not informed that the homophones indicated the corresponding colors, and they still focused on naming colors. The set of homophonous carriers was displayed in Table 3-22.

Table 3-22. Wordlist of sharing tonal syllable

Color Term	紅	棕	橙	黃	綠	藍	紫	灰	白	黑
IPA	xoŋ35	tsoŋ55	tʂʰəŋ35	xwɑŋ35	ly51	lan35	tsɨ21	xwej55	paj21	xej55
Meaning	Red	brown	orange	yellow	green	blue	purple	gray	white	black
Frequency	503	11	11	367	200	168	30	116	661	437
Onset Sharing	洪	鬃	懲	皇	慮	蘭	秄	輝	(百)	嘿
IPA	xoŋ35	tsoŋ55	tʂʰəŋ35	xwɑŋ35	ly51	lan35	tsɨ21	xwej55	paj21	xej55
Meaning	23	3	1	**	10	10	6	3	403	106
Frequency	洪	鬃	懲	皇	慮	蘭	秄	輝	(百)	嘿

** indicates there is no data or information

Most of the carriers were the same with the ones used in Homophonous Stroop's Task, so was their frequency correlation test. However, the lacking corresponding homophones of the color white "白 [paj35]" was substituted by "百 [paj21]" (hundred), which without sharing tone, in this shared unit section. It seems to be the last resort to use "百" to replace for color white, because there was no other bare syllable sharing words with compatible frequency. To keep the balance among all the shared unit tests, we still had to recruit "百 [paj21]" (hundred) to be the counterpart of "白 [paj35]" (white). Therefore, we used a character "百 [paj21]" in the unit sharing test, even though their tones do not share the same one. The result of the frequency correlation test between color terms and the homophones (syllable + tone) shows that the bi-tailed correlation between color terms and their corresponding carriers achieved the level of significance. The mean scores of the two variables are 2.50 and 38.11 respectively, while the scores of standard deviation are 228.86 and 62.96 respectively. The Pearson correlation is .834 ($p<.05$), which slightly reaches to the level of significance. It shows that the frequency distribution among the color terms still correlates the distribution among the tonal syllable-sharing carriers positively and significantly ($r=.834$, $p<.05$).

The trials in this word sharing were generally the same with former homophonous naming task. The first difference is that we put the carriers in a sequential way, rather than in pairs. The other is that we used “百 [paj21]” (hundred) to substituted for color “白 [paj35]” (white) in this section, while we didn’t recruit another characters to replace for white. In this tonal syllable-sharing test, we would like to put a visual difference between color term and the visual character to cause the linguistic competition, so we had to seek for another carrier to avoid the sound overlapping. In this section, we had 20 trials for subjects to name the color of the visual carriers.

3.2.7. General Procedure of Shared Unit Naming Task

Another 20 subjects participated in the experiment 3. Before this naming phase, subjects were asked to recognize the substituting characters. All of the subjects were checked that each pronunciation of the carriers can be recognized and uttered correctly, so the subjects were allowed to initiate the experiment. At the beginning of the testing phase, subjects could saw a star mark in the center of the screen for 2000 ms, which reminded subjects to focus on the coming trial. Then the visual stimulus was displayed, and subjects were asked to name the visual color for each of the carriers without time limit. During the answering period, the SONY-IC recorder was set aside to collect the answers which subjects provided. After finishing the naming, they had to push the response button on the serial response box to record the reaction time for single trial as well as to initiate next trial. When answering a trial, subjects were asked to look attentively at a single carrier and process them serially and one by one. It was prohibited to change, skip, reverse, or omit the processing order of carriers in a single trial. The above is a complete course for answering one trial. In shared unit naming task, there were six types of phonological unit tested, such as onset, vowel, rhyme, syllable, tone, and tonal syllable. Each of the target unit was designed 20 trials, which were recruited from the previous color naming task. Among the trials, 10 belonged to high group of phonological similarity, and the other trials were adopted from the low group. Therefore, for each of the subjects, they had to answer 20 trials for each of shared unit section, which means that there were 120 trials in total presented in those six sections of shared units. Twenty subjects were recruited in this experiment. That is to say, there were 2400 trials to be observed, and 19200 tokens (visual carriers) in total to be tested and analyzed.

After the experiment, all the sound files were transcribed, and the speech errors were detected and collected for later analysis. The criteria of error detection, record and categorization were the same with the previous experiment1 and 2. Three types of the relationship between target and error were also concerned in this shared unit test, such as phonological errors, semantic errors, and mixed errors. Regarding to the phonological and mixed errors (both errors are phonologically related), six types of phonologically structural relations were analyzed, such as onset, vowel, rhyme, syllable, tone, and tonal syllable. The phonological similarity of speech was also graded in this section.

E-prime served to help record the response time that subjects had done for trials. The response span among the six groups of shared united will be compared, and the temporal pattern would be compared to the pattern of speech errors for further discussion.

Chapter 4
Results and Discussions

From above five experiments, we have recruited 22 subjects in color naming, reading, Stroop naming, and homophonous naming test; another 20 subjects joined in shared unit test. As a result, we have collected 1,056 speech errors in total. In color naming test (Test 1), we have collected 96 speech errors; in color reading test (Test 2), there were 78 errors observed; in Stroop color naming test (Test 3), 257 speech errors were generated in this section. In homophonous naming test (Test4), there were 249 speech errors produced. In shared united test (Test 5), there were 376 speech errors being detected. The influence of independent factor, phonological similarity and phonological units, on the number of speech errors and response time will be discussed in the following sections.

The organization of this chapter appears as follows. The structure of lexical errors and reaction times among the tests will be compared and discussed in 4.1, and the role of phonological similarity and the modalities among color naming, reading, as well as Stroop naming tasks will be discussed in 4.1, too. Under independent factor of phonological similarity, we would like to examine the previous linguistic effects in chapter 2 and their relation to the number of speech errors and temporal data. As to the factor of phonological unit, such as initial, rhyme, syllable structure, tone, phonotactic constraint, will be analyzed and discussed in 4.2. As to the shared unit test, the issue on the possible units in lexical encoding will be examined and discussed in section 4.3.

4.1. The Structure of Speech error and Reaction Time: Task 1 ~ Task 4

In this study, 22 subjects participated in test 1 to test 4. There were 40 trials in respective tests, and 8 visual words were filled in each trial. In the following part, we will divide two portions: trial error frequency and error number. Trial error frequency, shown as "Trial F" in the following, indicates the error number that subjects made in those 40 trials of each test. The trial error number will be counted in high and low phonological similarity trials respectively. After data collecting, we can get several target-error pairs from each test, and we need to grade scores of phonological similarity for each pair. Then we can get the outcomes of high, medium, and low numbers in terms of phonological similarity, shown as "Error N" in the following. For example, in a trial from high phonological similarity group, subjects made four errors in this trial. The Trial F will be counted 4 in the high group and 0 in low group. Among these errors, we could get four target-error pairs. In terms of phonological criteria in table 3-2, one of them could be attributed to high phonological similarity group, another to medium group, and the others to low group. Therefore, in the column of Error N, we will mark 1 in high group, 1 in medium group, and 2 in low group. All the counts will be shown in percentage as well. Table 4-1 appears the structure of these speech errors in the four tests among the 22 subjects.

Table 4-1. The Structure of Speech Errors (N=680)

Speech Errors		Test 1		Test 2		Test 3		Test 4		Total
		Counts	%	Counts	%	Counts	%	Counts	%	
Trial F (N=680)	High	52	54.2%	52	66.7%	152	59.1%	118	47.4%	374
	Low	44	45.8%	26	33.3%	105	40.9%	131	52.6%	306
Error N (N=680)	High	37	38.5%	45	57.7%	150	58.4%	116	46.6%	348
	Medium	16	38.5%	12	15.4%	10	3.9%	17	6.8%	55
	Low	43	44.8%	21	26.9%	97	37.7%	116	46.6%	277
Phonological Effect		2.40		2.79		2.72		2.44		2.59

We have collected 680 errors among the four tests. With regard to Trial F, we observed that subjects produced 374 errors in the 20 trials of high phonological similarity, and 306 errors of low similarity. Except for test 4, the trials of high phonological similarity came up more speech errors (52:44 in test 1; 52:26 in test 2; 152:105 in test 3). Test 4 appears the opposite pattern. The trial of low phonological similarity brought out more speech errors (118:131 in test 4).

The phonological effect is the average score of the phonological similarity of the target-error pairs we collected in Error N. The average of the four naming and reading tests is 2.59. It would give us an anchor point that phonological effect for one of the tests to the average. In color naming test, the phonological effect is 2.40, 2.44 in homophonous naming test, 2.72 in Stroop naming test, and 2.79 in color reading test. It seems that the phonological effect weighs the heaviest in color reading task, and then Stroop naming task follows. On the other hand, it weighs the least in color naming test, and then the second least seems to be in homophonous naming test. It appears not only that phonological relation between target and error among the four tests might be different, but also that there might be different degrees of phonological dependency for respective visual tasks.

As to Error N, we graded the error pairs and divided them into three groups according to their phonological similarity. It seems that the pairs tend to be with high or low phonological similarity (625 in total), and there were only 55 pairs in between. Some samples are elicited in the following (1-4):

(1) xoŋ35 → xwɑŋ35

紅 (red) → 黃 (yellow)

Example (1) is classified as a lexical error with high phonological similarity. The criteria on grading the similarity is based on table 3-2. The score of phonological relation between *xong2* (red) and *xuang2* (yellow) is 5 points, since they share syllable number, syllable structure (CGVN in deep structure), initial [x], 35 tone, and nasal coda [ŋ]. This is a typical example of a speech error with semantic and high phonological relation in the case. There were 266 errors observed in this study.

(2) ly51 → xej55

綠 (green) → 黑 (black)

Example (2) is of a lexical substitution with less phonological relation. Between *lü*4 (green) and *xei1* (black), they only share the characteristic of syllable number. This case could be attributed to a pure lexical semantic error because color terms in Chinese could be used in monosyllabic form, as used in the experiments. Sharing syllable number seems to be an absolute result in this study. Therefore, we can categorize this kind of error as a pure semantic error. There are 187 errors in total which can be attributed to this case.

(3) xwɑŋ35 → pa35

黃 (yellow) → 拔 (to pull out)

Example (3) is attributed to the case of pure phonological error. Between the target and error, there is no semantic relation. As to the error unit between them, they only share tone, and the parts of syllable are substituted. There are just three errors involving syllable occur among the tests.

(4) xej55 → lan35

黑 (black) → 蘭 (orchid)

Example (4) is a case which shares no semantic relation, and less phonological similarity (except for syllable number). There are 173 errors of this case, but it is quiet unusual to have such high proportion (25.44%) of this semantic-phonological irrelevant case.

Test 2 and 3 shows that the number of pairs with high similarity exceeds the number of low similarity pairs (45:21 in test 2; 150:97 in test 3). In test 4, the numbers of the two groups tend to be equal (116:116 in test 4), while in test 1, the number of low similarity pair exceeds the number of high pair (37:43 in test 1). It is still vague to judge the phonological effect in error generating by error distribution.

From error distribution, it seems that the trails with high phonological similarity tended to induce more speech errors, and subjects tended to produce errors with phonological similarity. We need to put them under crosstab test to examine whether subjects have similar pattern in each test.

Table 4-2. Homogeneity of Proportions Among Subjects in Each Test

Chi-Square		N	x^2	df	Sig. (Pearson)
Test 1	Trial F	96	25.60	20	.179
	Error N	80	34.19	20	.025*
Test 2	Trial F	78	32.22	18	.021*
	Error N	66	30.66	18	.032*
Test 3	Trial F	257	52.42	21	.000**
	Error N	247	59.97	21	.000**
Test 4	Trial F	249	30.38	20	.064
	Error N	232	36.24	21	.021*

Note: *, ** are significant at the .05 and .01 levels respectively.

According to the chi-square test in table 4-2, we notice that Trial Fs in test 1 and test 4 pass the homogeneity test, which tells us that subjects have congruous pattern of error distribution when reacting to test trials. In test 1, after chi-square test, the error frequency of the trials among subjects doesn't reach at significance level (x^2=25.6, df=20, p>.05), while the Error N is significant at .05 level (x^2=34.19, df=20, p<.05). The result seems to be concordant with homophonous naming test. In test 4, it appears that Trial F among subjects doesn't reach at significance level (x^2=30.38, df=20, p>.05), but the Error N is significant at .05 level (x^2=36.24, df=21, p<.05). In square naming task and homophonous naming task, it seems that subjects might be sensitive to the phonological similarity of trials, but the target-error pairs didn't show congruous phonological distribution among subjects. In test 2 and 3, subjects showed a different pattern on Trial F and Error N. The Trial F and Error N in test 2 among subjects are significant (Trial F: x^2=32.22, df=18, p<.05; Error N: x^2=30.66, df=18, p<.05), and so as in test 3 (Trial F: x^2=52.42, df=21, p<.01; Error N: x^2=59.97, df=21, p<.05). The results show that, in color reading and Stroop naming tests, subjects' trial error frequencies were of difference as to phonological similarity of trials, and their speech errors didn't tend to appear similar pattern of high or low phonological similarity. It seems that phonological similarity did not cause apparent the same effect on each subject.

In order to know whether phonological similarity would cause subjects to react differently among the four tasks, we applied one-way ANOVA to examine the trial F distribution, as shown in table 4-3.

Table 4-3. One-Way ANOVA for Phonological Similarity and Trial Frequency

Trial Frequency	F ratio	Mean	df		SD	Levene		Sig.
			Between	Within		Levene	Sig.	
Test 1	14.76	2.18	3	172	1.88	11.13	0.000	.000**
Test 2		1.77			2.14			
Test 3		5.84			5.52			
Test 4		5.66			4.29			

Note: *, ** are significant at the .05 and .01 levels respectively.

With regard to the part of phonological similarity and Trial F, because the F-ratio (14.76) we computed exceeds the value of F (2.67) [4], we reject the null hypothesis and

[4] F value refers to the n_1 degrees of freedom (for greater mean square).

accept the scientific hypothesis that different phonological similarity differently affected the error frequency of trials that subjects made. It also shows that different phonological similarity would cause different Trial F among the four tasks [F(3, 172)=14.76, p<.01]. Table 4-4 brings out the post-hoc test (Scfeffe's post-hoc) and shows that test 1 (M=2.18, SD=1.88) and test 3 (M=5.84, SD=5.52) are significant at .01 level, and so do test 1 (M=2.18, SD=1.88) and test 4 (M=5.66, SD=4.29). Besides, test 2 (M=1.77, SD=2.14) and test 3 (M=5.84, SD=5.52) achieve at .01 level, and so do test 1 (M=2.18, SD=1.88) and test 4 (M=5.66, SD=4.29). However, the pairs of "test 1 x test 2" and "test 3 x test 4" don't reach to the .05 significant level.

Table 4-4. Post-hoc Analysis for Table 4-3 (Scfeffe)

Post-hoc Pairs	Test 1 Mean	Test 2 Mean	Test 3 Mean	Test 4 Mean	Sig.
Test 1 x Test 2	2.18	1.77			.968
Test 1 x Test 3	2.18		5.84		.000**
Test 1 x Test 4	2.18			5.66	.000**
Test 2 x Test 3		1.77	5.84		.000**
Test 2 x Test 4		1.77		5.66	.000**
Test 3 x Test 4			5.84	5.66	.997

Note: *, ** are significant at the .05 and .01 levels respectively.

According to one-way ANOVA statistics, it appears that Trial F distribution in naming task (test 1) differs from Stroop naming task (test 3), and it also differs from the pattern in homophonous naming task (test 4). Besides, the Trial F pattern in reading task also differs from the patterns in Stroop naming and homophonous naming tasks.

However, naming task didn't show difference from reading task (test 2), which implies that techniques of color naming and term reading would not cause different phonological sensitivity for subjects. Stroop naming task and homophonous task also show congruous pattern in their Trial F distribution. However, the above results imply that subjects always show different phonological sensitivity to deal with the trials from they were in naming and reading tasks when stimuli come with visual competition (test 3 and 4).

Since the phonological information of trials appears to affect error frequency for subjects, we further looked into the phonological relation of errors with their targets.

Table 4-5. One-Way ANOVA for Phonological Similarity and Error Number

Error Number	F ratio	Mean	df		SD	Levene		Sig.
			Between	Within		Levene	Sig.	
Test 1	15.56	1.82	3	172	1.80	12.395	.000	.000**
Test 2		1.50			1.98			
Test 3		5.61			5.49			
Test 4		5.27			4.14			

Note: *, ** are significant at the .05 and .01 levels respectively.

Table 4-5 was computed from one-way ANOVA test. Because the F-ratio (15.56) we computed exceeds the value of F (2.67), we could reject the null hypothesis and accept the scientific hypothesis that different phonological similarity caused phonological influence on the errors that subjects made. It also shows that different phonological similarity would induce different error numbers among the four tasks [F(3, 172)=15.56, p<.01]. Table 4-6 brings out the post-hoc test (Scfeffe's post-hoc). The result shows that the pair of test 1 (M=1.82, SD=1.80) and test 3 (M=5.61, SD=5.49) is significant at .01 level, and so does the pair of test 1 (M=1.82, SD=1.80) and test 4 (M=5.27, SD=4.14). Besides, the pair of test 2 (M=1.50, SD=1.98) and test 3 (M=5.61, SD=5.49) achieves at .01 level, and so does the pair of test 1 (M=1.82, SD=1.80) and test 4 (M=5.27, SD=4.14). However, the pairs of "test 1 x test 2" and "test 3 x test 4" don't reach to the .05 significant level.

Table 4-6. Post-hoc Analysis for Table 4-5 (Scfeffe)

Post-hoc Pairs	Test 1 Mean	Test 2 Mean	Test 3 Mean	Test 4 Mean	Sig.
Test 1 x Test 2	1.82	1.50			.983
Test 1 x Test 3	1.82		5.61		.000**
Test 1 x Test 4	1.82			5.27	.000**
Test 2 x Test 3		1.50	5.61		.000**
Test 2 x Test 4		1.50		5.27	.000**
Test 3 x Test 4			5.61	5.27	.979

Note: *, ** are significant at the .05 and .01 levels respectively.

Based on the result of one-way ANOVA, phonological similarity seems to induce different error distribution among tests. First, it seems that the error numbers in square naming test turned out to be different from Stroop naming and homophonous naming tests, but the error number distribution in Stroop naming test and homophonous naming test didn't show difference. Second, error numbers in reading test, Stroop naming test and homophonous naming test reached to the level of significant difference.

The computed result accords to the one of Trial F in table 4-3 and 4-4. Naming task didn't show much difference from reading task, and the pair of Stroop naming and homophonous naming tests did not, either. On the other hand, when visual competition comes out, the Error N appears different pattern from the one in square naming or term reading tasks. Based on the results of Trial F and Error N, phonological similarity appears to induce diverse patterns of error frequency and phonological relation between target and error when we cross-compared the four tests. Besides speech error, subject's reaction time (abbreviated as RT in the following) to each trial was also concerned in this study. The RT in individual tests was recorded and logged by E-prime experimental software. RT in this study refers to the time span that started from showing of single trial and ended in pushing the key to finish a trial. Therefore, the time span would include subject's answering, repetition, re-correcting, and halting during a trial. Table 4-7 is the one-way ANOVA to examine whether phonological similarity would affect subjects' react time among the four tests, which helps us know that phonological similarity could cause difference RT pattern in respective tests.

Table 4-7. Homogeneity of RTs Among Subjects in Each Test (one-way ANOVA)

RT	Test 1	Test 2	Test 3	Test 4	Average
High Group	5384.94	3724.34	6565.01	6183.85	5464.53
Low Group	5619.85	3559.03	6674.17	6419.07	5568.03
Average	5502.39	3641.69	6619.59	6301.46	5516.28

RT	F ratio	Mean		df		SD		Levene		Sig.
				Between	Within			Levene	Sig.	
Test 1	1.10	H	5.80	1	94	H	1237.38	1.95	.17*	.32
		L	6.06			L	1354.04			
Test 2	.49	H	3.72	1	42	H	777.22	.26	.61*	.50
		L	3.56			L	841.77			
Test 3	.11	H	6.57	1	42	H	1087.77	.01	.91*	.75
		L	6.67			L	1126.08			
Test 4	.42	H	6.18	1	42	H	1103.68	1.08	.31*	.52
		L	6.42			L	1304.52			

Note: *, ** are significant at the .05 and .01 levels respectively.

According to the result in table 4-7, we noticed that the F ratios among the tests (1.10 in test 1; .49 in test 2; .11 in test 3; .42 in test 4) do not exceed the F values (3.96 in test 1; 4.07 in test 2 ~ 4). Therefore, we could not reject the null hypothesis, and we cannot accept the scientific hypothesis that phonological similarity influence subject's reaction time in these tests, either. Furthermore, the values of significance in these tests do not achieve at .05 levels. It seems that, for subjects in each test, phonological similarity did not cause significant difference in RTs. On the other hand, the result seems to accord to the Stroop's work (1935) that term reading is processed faster than color naming.

Since we cannot infer that phonological similarity would cause effect in RT among subjects, we can see it still caused difference when we compared the four tests in pairs. The results are shown in table 4-8 and 4-9.

Table 4-8. One-Way ANOVA for Phonological Similarity and RT in Tests

<table>
<tr><th rowspan="2">RT</th><th rowspan="2">F ratio</th><th rowspan="2">Mean</th><th colspan="2">df</th><th rowspan="2">SD</th><th colspan="2">Levene</th><th rowspan="2">Sig.</th></tr>
<tr><th>Between</th><th>Within</th><th>Levene</th><th>Sig.</th></tr>
<tr><td>Test 1</td><td rowspan="4">71.13</td><td>5.47</td><td rowspan="4">3</td><td rowspan="4">172</td><td>1060.27</td><td rowspan="4">2.083</td><td rowspan="4">.104*</td><td rowspan="4">.000**</td></tr>
<tr><td>Test 2</td><td>3.64</td><td>805.01</td></tr>
<tr><td>Test 3</td><td>6.62</td><td>1095.54</td></tr>
<tr><td>Test 4</td><td>6.30</td><td>1200.06</td></tr>
</table>

After comparing the tests in pairs, we notice that the F radio is 71.13, which farther exceeds the F value 2.67, and it also pass Levene's homogeneity test. We could reject the null hypothesis and accept the scientific hypothesis that phonological similarity caused significant difference for the four tests [F(3, 172)=71.13, p<.01]. Table 4-9 provides the Scfeffe's post-hoc examination for the test pairs. If we compare the test with visual competition to another test without it, it appears that test 3 (M= 6.62, SD= 1095.54) and test 1 (M=5.47, SD=1060.27) are of significant difference, and so is the pair of test 3 and test 2 (M=3.64, SD=805.01). Besides, the pair of test 4 (M=6.30, SD=1200.06) and test 1 achieves at significant difference, and the pair of test 4 and test 2 does, too. Phonological similarity seems to induce different patterns of RT in Stroop naming and square naming tests, Stroop naming and term reading tests, homophonous naming and square naming tests, homophonous naming and term reading tests.

On the other hand, if we compare the pairs which are both without visual competition (test1 and 2) or with visual competition (test 3 and 4), we could get opposite results. Test 1 (M=5.47, SD=1060.27) and test 2 (M=3.64, SD=805.01) are significantly different, while test 3 (M=6.62, SD=1095.54) and 4 (M=6.30, SD=1200.06) do not show significant difference. That is to say that phonological similarity induced effect between naming test and reading test, but it did not induce significant effect between Stroop naming test and homophonous naming test.

Table 4-9. Post-hoc Analysis for Table 4-8 (Scfeffe)

Post-hoc Pairs	Test 1 Mean	Test 2 Mean	Test 3 Mean	Test 4 Mean	Sig.
Test 1 * Test 2	5.47	3.64			.000**
Test 1 * Test 3	5.47		6.62		.000**
Test 1 * Test 4	5.47			6.30	.004**
Test 2 * Test 3		3.64	6.62		.000**
Test 2 * Test 4		3.64		6.30	.000**
Test 3 * Test 4			6.62	6.30	.57

Based on the results of table 4-7 and 4-8, within individual test, phonological similarity could not induce significant difference in RT. If we compare RT data among the tests in pairs, it appears that phonological similarity caused significant difference according to the tasks which subjects took.

When we crossed-compare the results in table 4-2 (Trial F and Error N distribution) and table 4-7 (RT), we could merely see that phonological similarity in test 1 and 4 caused subjects to induce different distribution in Trial F, which indicates that subjects might have apparently different error frequency according to the phonological similarity of trials. The others in Trial F and Error N did not show up such significant difference, and RTs in the whole four tests did, neither. It seems the phonological effect did not affect extensively within single test.

On the other hand, if we crossed-compare the results in table 4-3 (Trial F distribution), table 4-5 (Error N distribution) and table 4-8 (RT), phonological similarity seems to induce significantly different distribution in Trial F, Error N, and RT among the four tests. Phonological similarity might cause above factors to act differently according to the visual task which was assigned to subjects. Later, we will have a discussion on the phonological effect within and between the visual tasks respectively, and on the relation from phonological effect to the factors of Trial F, Error N, and RT in the following section.

4.2. Phonological Effect on Lexical Encoding

4.2.1 Within Task

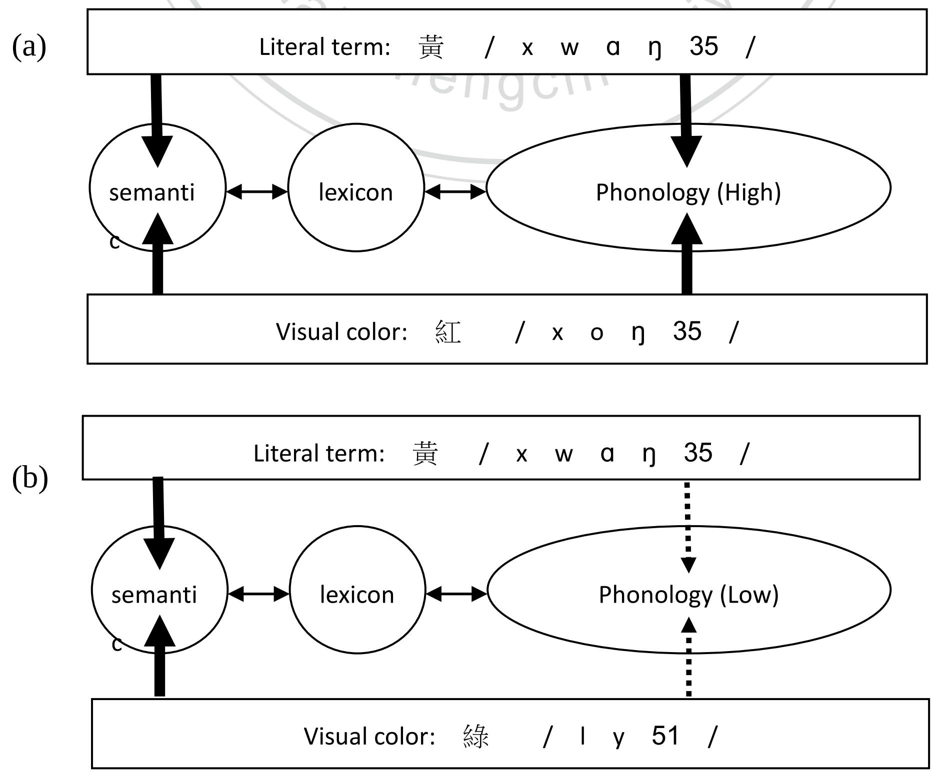
According to the computed results of table 4-2 and 4-7, except for the Trial F in test 1 and test 4, we noticed that the controlled factor of phonological similarity did not induce significant difference in Trial F, Error N, and RT distribution within respective test. With regard to Trial F, subjects showed phonological sensitivity to the trials in square naming test and homophonous naming test. In naming test, subjects tended to make speech errors to the trials with high phonological similarity; however, in homophonous naming test, they tended to make errors towards trials with less phonological similarity. They did not show apparently different error frequency in the other tests, so we could not say that phonological similarity brought out stable and apparent effect for subjects within each test. It seems that phonological similarity would not serve as a main effect for subjects to encode color terms when they were assigned to a certain visual task.

From the Trial F among the tests, we found that Stroop naming test induced the most speech errors (N=257), homophonous naming test follows (N=249) and term reading test was the least (N=78). It seems that test 3 and 4 induced apparently more speech errors than test 1 and 2. It appears that when visual color began to compete with literal information, subjects might produce more speech errors than just naming or reading. Even though phonological effect emerges only in test 1 and test 4, we still could not ensure whether subjects might induce more errors to the trials with more of less phonological similarity. With regard to Error N, we could not find out any tendency that subjects would produce errors with high phonological similarity because the computed statistics in table 4-2 did not show significant difference in Error N among the four tests. Phonological similarity did not induce a clear phonological effect when subjects produced speech errors. On the other hand, the statistics on RT within the four tests still could not show that trials with high or low phonological similarity would lead subjects to have apparent difference in processing speed. The above results reveal a fact that subjects would not have a consistent response and error patterns when phonological factor was controlled in certain task. It seems that phonological information might not be the only factor to be processed during lexical encoding.

4.2.2 Between Tasks

If we take the four tasks into consideration at the same time, according to the computed results in table 4-3, 4-5, and 4-8, we could see that, the controlled factor caused subjects to reacted differently to trials with phonological similarity (Trial F and RT), as well as their error production (Error N) when we cross-compared the four tasks. There would be different phonological dependency according to the visual task to which they were assigned.

With regard to the Trial F, Error N, and RT among the four tasks, phonological factor seems to induce effect when subjects were under different visual tasks. Even though there is less phonological effect within each test, the effect appears more when the four tasks are compared. If we compared either naming or reading test to Stroop naming or homophonous naming test, subjects showed different Trial F distribution between each pair. The result implies that if the literal information interfered in naming mechanism, subjects showed more phonological effect than just naming or reading terms. The possible reason could be that naming or reading reflects more close to facts in our lexical process in life, the effects from individual linguistic levels could be balanced. The effects from linguistic levels would still be balanced in Stroop naming and homophonous naming, but their entire strength of phonological effect in lexical network might be higher than the strength in naming or reading task, which could serve as a reason to explain why phonological effects exists when we crossed-compared the results for these tasks. With a view to the network strength, as shown in figure 2-6 (Dell & O'Seaghdha, 1991), connectionism provides an extended explanation that Stroop technique brings more than one extra signals, literal signal and visual color signal here, into the lexical network. The simulated process of such lexical network is depicted in figure 4-1.



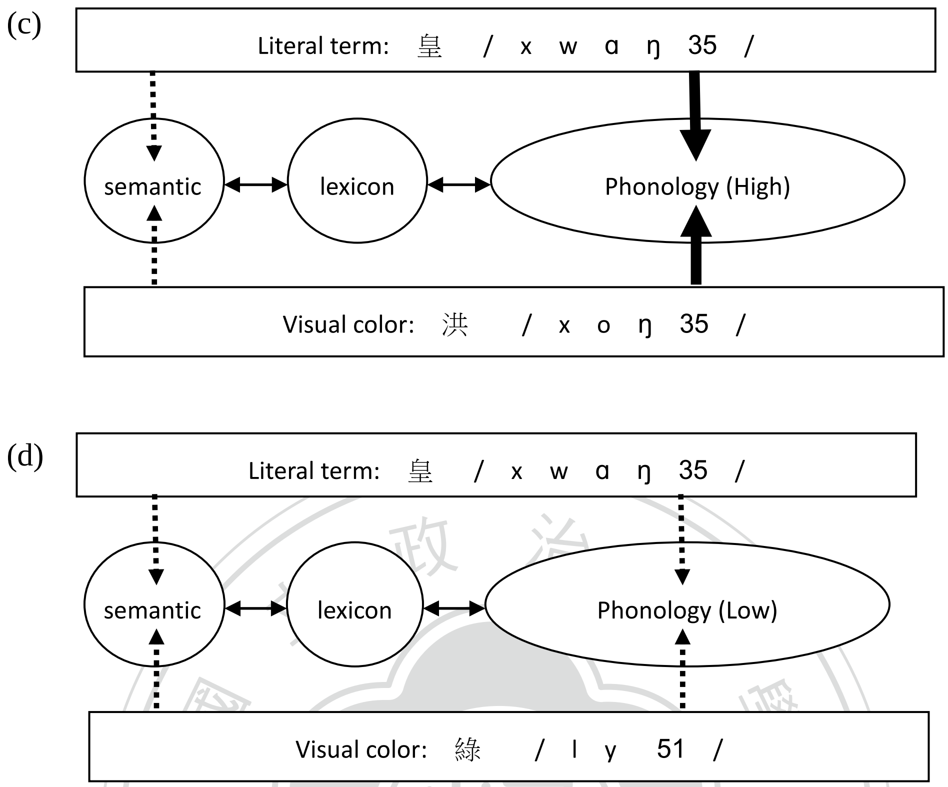


Figure 4-1. Simulated lexical network for Stroop technique

Compared with naming or reading process in figure 2-6, the simulated processes in figure 4-1 were provided to depict the activation strength in the lexical network when subjects were assigned to take task 3 and 4. The process (a) and (b) indicate the Stroop naming task, with high or low phonological similarity in trials respectively, while (c) and (d) simulated the process of homophonous naming task. The entire network strength is the greatest in (a). The visual color and term are both color concepts, so they got dual semantic activations from the two channels. Even though subjects were told to name the colors, the activation from literal term still existed, which could wrongly attract subjects to read the term and produced speech errors. Process (b) shows if activations come from dual visual sources with low phonological similarity, the strength of phonological layer would get decreased, and the strength of semantic retained for both of them were colors. The strength of whole network became less than (a). Homophonous naming test reduced the strength in semantic layer because the literal term and visual color did not belong to the same semantic field. However, the strength of activation in phonological layer retained strong because of high phonological similarity between the dual concepts. The strength in phonological layer as well as in semantic layer might get declined when the dual concepts showed little phonological similarity, as depicted in process (d).

According to above statistic results, process (a) and (b) did not show significant difference in network strength, and process (c) and (d) did not, either. It seems that reduced strength of phonological layer would not be significant, so there was nearly little phonological effect within each visual task. However, if we compared the network in naming or reading task to Stroop naming or homophonous naming task, the strength in network appears significant difference. The effect of phonological strength might get greater when visual competition came up, as in (a), (b), (c), and (d). The dual activations through visual channel at the same time caused phonological effect to be greater than task 1 and 2, and it also led subjects to produce rather more speech errors in task 3 and 4. It could help us explain why phonological effect induced significant difference in Trial F, Error N, and RT between tasks. The extra strength from more than one input could cause entire activation strength to be more enhanced and confused for subjects than the strength from single input.

4.2.3 Generation of Speech Errors and RT in Stroop's Tasks

Connectionist model provides a probability for us to explain how retrieving errors came out by means of Stroop technique. Dell & O'Seaghdha's model (1991) depicted the way which external signals affect the lexical process, as in Figure 2-6. In their model, external signals could come in from any layer of lexical process, such as from semantic, word, and phonological layers at the same time, and relevant nodes in each linguist representation were ready to be activated. Even though subjects were asked to focus on specific visual target (color), the literal term also sent external signals into linguistic layers with the visual color information meanwhile. The mental process is supposed to be "busier" because subjects should resist the strength of activation from color term, which is called the "noise" of activation in the lexical process. The coming out of Stroop errors might be the wrongly activated process and encoded to the motor generator. The busyness of the lexical network could also explain why it took much more time for subjects to process trials in task 3 and 4. The external signals from more than one source coming at the same time could explain the longer time that subjects spent in processing Stroop technique.

4.3. Linguistic Effects and Speech Errors: Test 1 ~ Test 4

In this section, we will display and discuss the linguistic effects among the color naming, color reading, Stroop naming, and homophonous naming tests. These effects include initialness effect, rhyme effect, tone effect, phonotactic regularity effect, and Stroop effect. Table 4-10 shows the structure of these linguistic effects, and then the following sections will discuss these effects during lexical process. The total number of speech errors here is 680 (N=680). The following table displays the phonological distribution of these linguistic effects among all target-error pairs we collected in the four tasks. Table 4-11 is the computed result of one-way categorical ANOVA, and its post-hoc analysis will also be shown in table 4.12.

Table 4-10. The Counts of Linguistic Effects among All Errors (N=680)

Linguistic Effects	Task 1		Task 2		Task 3		Task 4		Total	
	Count	%	Count	%	Count	%	Count	%	Count	%
Initial	38	39.6%	42	53.9%	94	36.6%	75	30.1%	249	40.0%
Rhyme	8	8.3%	14	18.0%	25	9.7%	17	6.8%	64	10.7%
Prenuclear	5	5.2%	9	11.5%	52	20.2%	28	11.2%	94	12.1%
Vowel	0	0.0%	2	2.6%	5	2.0%	22	8.8%	29	3.3%
Coda	14	14.9%	10	12.8%	41	16.0%	31	12.5%	96	14.0%
Structure	36	37.5%	35	44.9%	137	53.3%	110	44.2%	318	45.0%
Tone	43	44.8%	33	42.3%	113	44.0%	103	41.4%	292	43.1%
Phonotactic	96	100.0%	78	100.0%	257	100.0%	249	100.0%	680	100.0%
Stroop	No Data		No Data		221	86.0%	189	75.9%	410	81.0%

Table 4-10 provides the result of counts after we analyzed and categorized for all of the target-error pairs. The data shows the numbers and proportions of these units that the errors share with their targets. Phonotactic effect is an overwhelming effect that all the errors follow in all tests, which means subjects never produced a word whose phonological structure did not exist in their language. Stroop effect seems to be the secondary effect to affect when there were dual visual representations coming up at the same time. We need to put above data into statistic examination for further observation and discussion, as shown in the following.

Table 4-11. One-way ANOVA Results of Speech Errors

One-way ANOVA	F value	Mean	df		SD	Levene		Sig.
			Between	Within		Levene	Sig.	
Initial	8.22	62.25	8	25	26.89	44.65	.000	.000**
Rhyme		16.00			7.07			
Prenuclear		23.50			21.49			
Vowel		7.25			10.05			
Coda		24.00			14.54			
Structure		79.50			51.99			
Tone		73.00			40.83			
Phonotactic		170.00			96.18			
Stroop		205.00			22.63			
Total		65.65			12.20			

Table 4-12. Post-hoc Analysis for Table 4-11 (Scfeffe)

Post-hoc Pairs	Rhyme	Prenuclear	Vowel	Coda	Structure	Tone	Phonotactic	Stroop
Initial	.96	.99	.90	.99	1.00	1.00	.18	.12
Rhyme		1.00	1.00	1.00	.81	.89	.01**	.01**
Prenuclear			1.00	1.00	.89	.95	.02*	.02*
Vowel				1.00	.68	.78	.01**	.01**
Coda					.90	.95	.02*	.02*
Structure						1.00	.39	.23
Tone							.30	.18
Phonotactic								1.00

Note: *, ** are significant at the .05 and .01 levels respectively.

According to the computed result in table 4-11, because the F-ratio (8.22) exceeds the value of F (2.34), we could accept the scientific hypothesis that the counts of these linguistic effects show significant difference [F(8,25)=8.22, p<.01]. It appears that these linguistic effects were found to impact error generation differently. The total mean value among the four tests which one-way ANOVA generated is 65.65, which provides a basic level to justify what kinds of effects influenced and dominated the generation of speech errors. It seems that the mean values of tone, structure, phonotactic, and Stroop effects exceed the total mean value. These effects are more prominent than the other linguistic effects, such as the effects of vowel, rhyme, prenuclear glide, coda, and initial. However, the mean value could not help declare certain of these effects exist in generation of speech errors or in lexical encoding. Since we know these effects appear difference among errors, we need to know how each effect have contract with all the other effects. These effects were compared to see their difference by means of Scfeffe post-hoc test in table 4-12, and we will go over all of these linguistic effects in the following and have discussion on their individual effects in lexical network.

4.3.1. Initialness Effect

According to table 4-10, the total amount of the errors sharing the initial part with the target is 249, and the occurring frequency among speech errors is 40.03%. Color reading task induced the most errors which share initial part, whose frequency reaches to 53.85%. There were about half of errors in reading test which tended to preserve the initial part, or to retrieve a lexicon which shares the initial. Naming test, Stroop test, and homophonous test shows that the rate of onset-sharing is between 30% and 40%. As to proportion, it seems to be consistent with the findings of Dell (1986) that initial is always detectable and salient in phonological structure. With the account of interactive process model, the salient structure which is already activated in phonological layer also sends feedback to the nodes of lemma layers, and sometimes retrieves the inaccurate lexicon with the same initial, as well as the relevant lexical meaning. However, according to the result of one-way ANOVA in table 4-11 and 4-12, we noticed that not only that the mean of initialness (mean=62.25) did not exceed the total mean 65.65, but it did not show any significant difference with the other effects in the post-hoc result. It means that the errors sharing initial with their targets did not achieve to a statistically salient amount. The effect of initialness might not be as apparent as Dell (1986) mentioned in this study.

As to the phonological similarity, we see that speech errors with high phonological similarity still rely greatly on the information of syllable initial. The counts of initialness override the amount of rhyme, vowel, prenuclear glide, and coda individually. It implies that syllable initial is a rather salient structure among all phonological structures in lexical process, which not only helps us retrieve target lexicons correctly, but also retrieve the wrong lexicons with the same onset backwards from phonological layer. However, the issue of whether initial is a facilitative or an interfering effect in lexical will be examined and discussed in the section of advance planning unit. At this phase, we could only assume and infer that initialness would not induce a salient effect, but induce certain amount, when speech errors are generated.

4.3.2. Rhyme Effect

In the four tasks, the proportion of the errors preserving rhyme comes to 10.71% (N=680, Rhyme Preserving=64). Comparing rhyme effect to initialness effect, the percentage of speech errors with rhyme sharing (10.71%) is far less than the percentage of sharing syllable onset (40.03%). According to the post-hoc result, we found that the count of rhyme only contrast with the amounts in phonotactics ($p=.01$) and Stroop effect ($p=.01$), but there is no significant difference between rhyme and any other sub-syllable units, such as initial, vowel, prenuclear glide and tone ($p>.05$). Apparently, rhyme effect seems not to impose influence as greatly as initialness. Rhyme could not be a salient phonological organization which could lead to retrieve a lexicon sharing the same content of rhyme, especially in such a strong lexical network of colors.

The issue of content (segments) in rhyme structure seems not to induce abundant rhyme sharing errors. The subordinate contents, prenuclear glide, vowel, and coda, did not show any significant difference in their counts with other possible effects, but they contrasted with the numbers of phonotactics and Stroop effect significantly, as rhyme did. Therefore, we could claim that rhyme, including its subordinate contents, did not show any prominent effect in error generation or lexical encoding. On the other hand, the effect of structure in rhyme will be involved in the section of phonological structure.

4.3.3. Tone Effect

The number of the errors sharing tone with targets is 292 in total, which occupies 43.11% of all. The mean value of tone is 73.00, which exceed the value of total mean 65.65. Tone seems to act more prominent than initial with regard to their numbers and mean values. According to the post-hoc result, the number of tone did not show apparent difference with other linguistic effects, as the same with initialness. It still hang us a vague area for us to judge whether effect play a dominant role in lexical encoding. One thing for certain is that tone effect weighed over the initialness effect in this study. If the tones in Chinese have equal chance to substitute for each other, the chance estimate of tone replacement is supposed to be 25%. Then the percentage of 43.11% seems to imply that tone effect affects and dominates the generation of lexical substitution. We could not deny that this effect exists, but we noticed that tone effect is rather more significant than initial effect and rhyme effect.

The distribution of error counts seems to tell us that tone might be different from the structure of phonemes, such as initial and rhyme. The proportion tells us that initial (53.9%) and rhyme (18%) appear to affect greater on autonomous lexical process (reading task) than the other tasks, but tone effect appears nearly fair among the four tasks, within 41% to 45%. It seems that tone is more than a pure phonological structure, and it is supposed to be attributed to a larger framework, such as a phonological organization in lexical structure.

Among the four tests, we could not see any lexical substitution with pure tone substituting, omitting or addition. All of the substitutions mapped to colors, except for one tone-sharing case of “*xuang* [xuɑŋ35] “ (yellow) which is replaced for “*ba* [pa35] (pull)” and could not map to any colors. The exception could not be the evidence of independent status of tone in Chinese. Since the fact that target tone tends to affect the lexical retrieval within the same semantic domain, it seems to support the viewpoint of tone status in the works of Wan & Jaeger (1998) and Wan (2007) that tones are represented lexical underlyingly and ought to be part of the phonological organization of the lexicon, rather than the view of phonological frame acted like stress in English, which was proposed by Chen (1999).

4.3.4. Syllable Structure Effect

The amount of errors sharing phonological structure with targets is 318, which occupies 45% among all errors. The proportion within individual task is from 37.5% to 53.3%. The mean value of stricture is 79.5, which exceeds the total mean value 65.65. According to the post-hoc result, there is no significant difference in number with other possible effects. It means that syllable structure is not necessary information in lexical encoding, but a salient effect. From the results of proportion and mean value in one-way ANOVA, it still appears that syllable structure effect affects and dominates the error generation and lexical encoding.

In order to explain the generation of errors with the same phonological with targets, as discussed in former section, the selected phonological frame and relevant nodes would send activating weight backwards to the lexicon layer.

When a color is processed, the lexical nodes with the same phonological frame as well as relevant meaning might get ready to be activated at the same time. Then, the imprecise feedback from phonological layer might have higher probability to retrieve the lexical nodes with the same phonological frame, which causes the lexical substitution to occur more frequently with the same phonological structure.

When we look into the percentage of syllable sharing of these speech errors, we found that the effect of syllable structure affected differently among the four tests. In naming task, 96 errors were generated, there were 47 cases occurred with syllable sharing, which occupied 48.96% in task 1. As to reading task, there were 78 errors in total, and 42 cases were produced to share syllable structure with targets, which took 53.85%. With regard to the Stroop naming task, there were 257 errors in total, and 171 errors, which occupied 66.54%, shared phonological similarity with target color. In homophonous naming task, there were 249 errors in total, and 143 errors shared phonological structured, which occupied 57.43%. Except for task 1, the other tasks induced more than half errors with the same syllable structure of targets. The computed statistic result of Chi-square test shows that the cases of syllable structure sharing among the four tasks appears significant difference ($x^2=11.15$, df=3, $p<.05$). Therefore, we could claim that Stroop naming task seems to induce a stronger effect on syllable structure than all the other tasks, and homophonous naming task follows. The possible reason might be that Stroop naming involves processing visual color and color term at the same time, and the dual input might lead to a better preparation for phonological framework in this task before lexical activation. Therefore, the error would be easily to be encoded with the same syllable structure. Even though there is no semantic relation between the dual inputs in task 4, phonological structure still affected the generation of speech errors. In addition, in comparison of task 3 and 4, we noticed that the effect was quite stronger when dual inputs were in semantic relation, as in task 3.

Generally speaking, syllable structure effect apparently exists in lexical process, especially in the case of Stroop technique.

4.3.5. Phonotactic Regularity Effect

According to the result in table 4-10, none of phonotactic errors were found in this study. It shows that phonotactic regularity is an absolute effect which affects and dominates the lexical process. There were no speech errors which violate the phonological constraint in Mandarin. Dell (1993) proposed that there should be phonological frame for legitimate sound sequence in phonological representation to be checked and encoded. To explain this phenomenon, phonological frame might act like a syllable frame (schema), and only the legal sequence of sounds could be encoded in an available frame. Since there are no illegitimate frames to map, little phonological violation could be encoded to phonetic level and produced. After a lemma is activated in lexical layer, certain phonological frame and relevant phonological nodes in the phonological layer could be activated meanwhile. That could be the reason why there were not any phonotactic errors to be generated in these experiments. A lexical network which is prepared to be activated might let the relevant lexical and phonological nodes get ready before activation. High phonological similarity might lead to a neighboring lexical node to be selected (lexical substitution), but phonological frame could help prevent the activation at lemma level from forming phonotactic errors during phonological encoding.

4.3.6. Stroop Effect

By means of Stroop technique, we could induce the errors involving visual representation. We, tentatively, name this type of error after the pioneer's name—"Stroop error" in this study. In table 4-10, there were 86% of speech errors produced by reading the term rather than naming the color in task 3, and there were 75.9% of errors by the same way in task 4. The mean value of this effect is 205, which far exceeds the total mean value 65.65, so this effect seems to be obvious and dominant in lexical encoding. According to the post-hoc result, we found that number of Stroop type contrasted significantly with the numbers of rhyme, prenuclear glide, vowel, and coda.

In order to account for the Stroop errors, we would like to analyze the lexical process in terms of connectionist model in figure 4-1. In task 3, when the trials presented, the dual inputs caused the relevant nodes in lexical and phonological layers to be prepared for activation. It seems to be reasonable that errors came from reading color terms by accident during naming task, because the wrong visual representation induced relevant nodes being activated and encoded to the motor program. These nodes would receive more strength than the nodes beyond the dual inputs. That could serve to account for why the number of Stroop errors always exceeded producing other error naming. If the dual inputs were both attributed to color lexicon, as in task 3, the chance to retrieve the literal term might get a little higher. If there were no semantic relation, as in task 4, the chance might get decreased slightly. The strength of the whole network still played a crucial part in generating speech errors and the chance of Stroop error.

In current study, the classical Stroop technique opens up another window for lexical processing and linguistics issue. Stroop errors provide a new direction to explain the generation of speech errors resulting from visual representation. We always focus on one visual task during lexical processing, but other inputs which is not intended also come into our sensory channel and make relevant nodes in lexical network ready to be activated. If there are more similarities in their linguistic characteristics, and the strength of lexical network for dual inputs would be stronger. Therefore, the chance of retrieving the wrong lexicon might get higher.

To sum up, table 4-13 is shown to conclude the above linguistic effects we have discussed so far.

Table 4-13. Summary of Linguistic Effects

Effects	Saliency
Initialness	Salient in number
Rhyme	No
Vowel	No
Tone	Salient
Phonotactic Regularity	Salient
Syllable Structure	Salient
Stroop Effect	Salient

4.4. The Structure of Speech error and Reaction Time: Task 5

In order to testify the possible units in lexical process, we conducted shared unit test and observed their processing speed and error amounts. In this experiment, we recruited six sets of carriers which share different phonological units with the target colors, and we asked subjects to name the visual colors instead of reading the characters. Table 4-14 and figure 4-2 show the results of speech errors and response time in shared unit test.

Table 4-14. The Structure of RT and Errors N in Test 5 (N=376)

Shared Units	Onset	Vowel	Rhyme	Syllable	Tone	Syl.+tone	Average
RT	5446.30	5435.59	5264.84	4970.58	5470.20	4571.08	5193.09
Errors N	70	78	69	57	65	37	62.7
Percentage	18.62%	20.74%	18.35%	15.16%	17.29%	9.84%	376

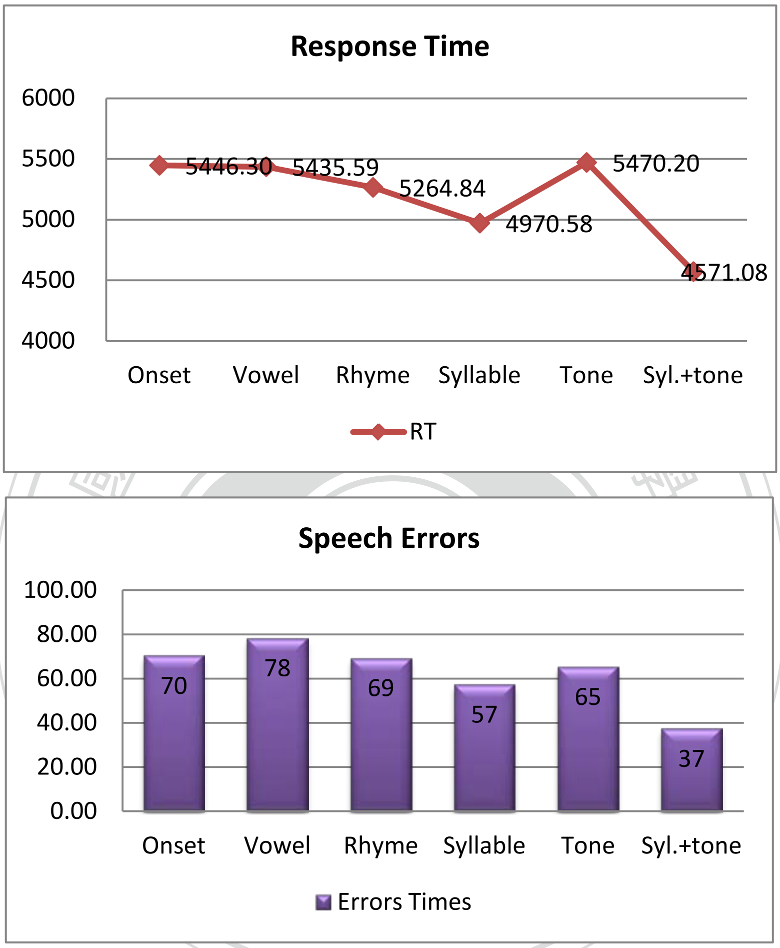


Figure 4-2. Response Time and Speech Error in Experiment 3

From the reaction time among the units, tone-sharing unit took the most time for subjects to name the colors (5470.20ms). Onset-sharing unit was similar with vowel-sharing unit (5446.30 and 5435.59 ms respectively). The response time in rhyme-sharing unit was faster, the average span is 5264.84 ms. When subjects reacted to the syllable-sharing trials, the response span fell to 4970.58 ms. The trials whose color and term was tonal syllable-sharing were processed the fastest among these units, which could fasten to 4571.08 ms. In order to examine the target units induced a facilitation or inhabitation effect in lexical process, we need the statistics to testify these possible units.

Table 4-15. One-way ANOVA for Error Counts Among Target Units

Error N	F value	Mean	Df		SD	Levene		Sig.
			Between	Within		Levene	Sig.	
Onset	2.35	5.15	5	288	3.92	12.51	.000	.041*
Vowel		4.49			2.14			
Rhyme		4.36			2.23			
Syllable		4.14			1.55			
Tone		4.17			2.29			
Syl.+tone		3.28			1.67			
Total		4.35			2.50			

According to the computed result, the F-ratio (2.35) exceeds f value (2.26), and we could accept the scientific hypothesis that these phonological units could induce significant difference in error amounts [F (5, 172) = 2.35, p<.05]. The total mean value is 4.35, the mean of these units exceed 4.35 is onset, vowel, and rhyme, which indicates that these units induced rather more speech errors. The other units, such as syllable, bare tone, and tonal syllable, induced less speech errors in this test. We need to refer these results to the post-hoc test in table 4-16.

Table 4-16. Post-hoc Analysis for table 4-15 (Scfeffe)

Post-hoc Pairs	2. Vowel	3. Rhyme	4. Syllable	5. Tone	6. Syl.+tone
1. Onset	.74	.54	.30	.35	.02*
2. Vowel		1.00	.98	.99	.27
3. Rhyme			1.00	1.00	.39
4. Syllable				1.00	.67
5. Tone					.65

Note: *, ** are significant at the .05 and .01 levels respectively.

According to the results of post-hoc in table 4-16, we noticed that only the pair of onset and tonal syllable were significant different in their error numbers. The other pairs were of no significant difference. As to the results of table 4-15 and 4-16, the unit of initial caused the most speech errors among subjects, and the unit of whole tonal syllable induced the least. Their amounts of speech error achieve to significantly different level.

Generally speaking, if we compare the amount of speech error among these possible units, it appears that phonological unit really affects the error amount significantly. From the point of error amount, we found that tonal syllable could be a facilitative unit in lexical process, while the unit of initial might be an effect of inhabitation. From the results of 4-15 and 4-16, it seems that phonological units affect the process of lexicon, especially when external competition comes in visual channel. We need to compare the result of error amount to the one of response time in this test.

Table 4-17. One-way ANOVA for RTs Among Target Units

Error N	F value	Mean	Df		SD	Levene		Sig.
			Between	Within		Levene	Sig.	
Onset		5.60			876.89			
Vowel		5.59			927.80			
Rhyme	1.43	5.38	5	623165	833.45	187.90	.000	.000**
Syllable		5.13			940.75			
Tone		5.61			922.10			
Syl.+tone		4.74			934.07			
Total	F < 2.21	5.36			956.38			

According to 4-17, weighted by the factor of shared phonological units, the data of response time shows the value of F value is 1.43, and it reaches to the significance level. We could reject the null hypothesis and accept the scientific hypothesis that different phonological units caused significant difference of reaction time [F(5, 623165)=1.43, p<.05]. Therefore, the result seems to support the research assumption that phonological units might cause facilitation effect or inhibition effect to occur in lexical process. Compared with the result in table 4-14, we could say that the tonal syllable-sharing and bare-syllable sharing units could induce the most facilitation in naming task, while bare tone could not. Table 4-18 shows the post-hoc analysis among these target units.

Table 4-18. Post-hoc Analysis for table 4-17 (Scfeffe)

Post-hoc Pairs	2. Vowel	3. Rhyme	4. Syllable	5. Tone	6. Syl.+tone
1. Onset	.85	.000**	.000**	.001**	.000**
2. Vowel		.000**	.000**	.000**	.000**
3. Rhyme			.000**	.000**	.000**
4. Syllable				.000**	.000**
5. Tone					.000**

Note: *, ** are significant at the .05 and .01 levels respectively.

With the computed result in table 4-18, the difference seems to be more significant than that in error number. Except for the pair of onset and vowel, all of the other pairs appear significant difference in response time. Therefore, we could provide a hierarchy of response time (from fast to slow) among these phonological units, as shown in (1). In addition, we also provide the hierarchy of error number (from few to many) among these units in (2), based on the result in table 4-14 and 4-15.

(1) RT Cost Hierarchy of Phonological Units:

Tonal syllable < syllable < rhyme < vowel< onset < tone.

(2) Error Number Hierarchy of Phonological Units:

Tonal syllable < syllable < tone < rhyme < vowel < onset.

Tonal syllable is the unit which subjects processed the fastest and induced the fewest errors than all the other units. We could state that tonal syllable induced the most facilitation effect in lexical process. Syllable could be the secondly fastest unit for subjects to process lexicon as well as the unit to produce penultimatly more speech errors among these units. From the statistics above, we found that the mean values of error number and RT were both below the total mean values, and they passed the significant level of one-way ANOVA in table 4-15 and 4-17. The two units induced rather fewer speech errors and processed faster than the other phonological units. We could say that the units of tonal syllable and syllable sharing could be a facilitative effect in lexical encoding.

Rhyme is the thirdly facilitative unit for subjects to encode a lexicon when dual input came in visual channel, but it induced speech errors the third more from the last, which is more than the unit of tone. It seems to be a watershed among these units. In table 4-15 and 4-17, not only the mean values both exceed individual total mean value, but both of them passed the statistic examination. It indicated that the rhyme sharing unit induced more speech errors and more processing time than the average, which could help to judge that rhyme could be an effect of inhabitation, and rhyme-sharing could not help subjects to reduce speech errors significantly.

The unit of vowel was processed slower than the unit of rhyme and it also induced more speech errors than rhyme-sharing trials. The unit of onset was processed the slowest, and it affected subjects to produce rather more speech errors, which is only fewer than the unit of tone. The mean values of vowel and onset do not exceed their total mean values, and it passed the examination of statistics as well, as shown in table 4-15 and 4-17. The results reveal that onset or vowel sharing units could be an effect of inhabitation in lexical process.

The unit of tone seems to be a special status in lexical process. In these hierarchies, tone produced the thirdly fewer speech errors, but it was the most halting in RT among all units. It is hard to judge whether tone is a facilitative unit or not from these hierarchies. However, according to above statistic results in 4-15 and 4-17, we found the mean value in error count went below the total mean value, but it exceeded the total mean value of RT the most. It shows that the unit of tone is effective in reducing the number of speech error, but a kind of inhabitation in processing speed. If we regard the hierarchy in (1) and (2) as a continuum, the left side is lexicon-like, and the right side is segment-like. Tone could be attributed to a segment-like unit in error amount, but it should be regarded as a lexicon-like unit in response time. Tone did not have any significant facilitation in speed when subjects processed tone-sharing in trials, but it could generate more precise and correct lexicon than the units of rhyme, vowel, and onset. Even though tone did not affect as great as the unit of syllable, it still appears a certain of lexical quality. The possible account is that tone might be a lexical organization and encoded in phonological representation, not a pure phonological tone. If it is a phonological tone, the patterns of error number should act as the phonological units of onset, vowel, and rhyme. Therefore, the status of tone seems to support the proposal of Wan & Jaeger (1998) and Wan (2007) that tone is a phonological organization of a lexicon. The following table is the summary among these phonological units.

Table 4-19. Summary of Phonological Units

Units	Criteria	Effect
Onset	Error N	Inhabitation
	RT	Inhabitation
Vowel	Error N	Inhabitation
	RT	Inhabitation
Rhyme	Error N	Inhabitation
	RT	Inhabitation
Bare Tone	Error N	Facilitation
	RT	Inhabitation
Syllable	Error N	Facilitation
	RT	Facilitation
Tonal Syllable	Error N	Facilitation
	RT	Facilitation

4.5. Summary: Reflections on Lexical Process Models

The traditional serial account, such as in Fromkin (1971), single input is the central issue on lexical process. The input could go on to next linguistic section after is processed in former linguistic generator. If the process brings out some problems at certain stage, speech errors will be generated. Serial model provides a reasonable space for linguistic rules to be operated in individual stages, even if the speech errors are generated. However, Fromkin's model is not sufficient enough to explain the errors form dual external inputs and the way that speech errors are retrieved. The input of each linguistic department is always from single source, including the first stage of meaning to be conveyed.

Levelt's model (1989, 1999) provided serial model a better explanation for the way that speakers retrieved speech errors, by means of the nodes, links, activation and spreading within lemma level. The neural-linguistic approach provided a psychological account for the wrong retrieving. In order to explain the lexical error with phonological similarity, Levelt's model explained that nodes with more similarities could have higher chance to be altered and activated. Levelt's model provided a possible basis for accounting for the dual inputs, such as Stroop technique in this study, because the external inputs could have competition after they are activated and spread. The competition could tell us that the target color or speech error would be generated in the end. However, the model could not have enough explanation for why color term and homophone could induce significant difference in error distribution and RT. It is difficult to explain the competition of semantic and phonological levels occurred in task 3 and 4 without interaction between linguistic layers, as depicted in figure 4-1.

With a view to the interactive models, Stemberger (1985) proposed that the linguistic departments should be interactive, and the activation is transported through units and links. It still lacked sufficient account for the operation of dual inputs during lexical process.

Connectionist model, especially the model of Dell & O'Seaghdha (1991), provided enough explanation for dual external inputs. The model also accounted for the interaction of semantic, lexicon and phonological layers, as shown in figure 4-1, and the way how task 3 and 4 induced difference in error amount and response time. Based on the results in linguistic effects and shared unit test, this model still provided limited space for the reason why certain units induced facilitative effect (tonal syllable, and syllable), and why errors always share phonological structure and tone with visual targets. However, this model really helps explain how Stroop effect operates in lexical network and how it generates speech errors.

It shows that when a certain lexical domain (color in this study) is activated, and all the lexicon and phonological lines and nodes are prepared to be activated, which forms a strong network for further lexical processing. Therefore, it is scarce to see any error occurring without meaningful or phonological relation in this study. Beside activation of forward spreading (cascading from lexicon to lexeme layer), backward spreading (feedback from lexeme to lexicon layer) also works to retrieve sound-similar color in error generation. With the interactive account, discrete departments could communicate among layers, so that prominent phonological information is traceable in color substitution errors. Therefore, we could see that both semantic and phonological effects are equally crucial in visual tasks. Phonological effect suggests that when the semantic domain is specified and gets ready for activation, all the network of relevant semantic nodes, lexicon nodes, and phonological nodes would get ready at the same time (or maybe gradually). These relevant nodes form a stronger network before specific activation, and relevant semantic and phonological nodes are sensitive to be activated. By means of observing the linguistic effects in speech errors and the shared unit test, it appeared that dual visual inputs could have co-effect in generating speech errors and encoding speed, especially within the network between the semantic and phonological layers. The hierarchies in (1) and (2) show the hierarchies of phonological sensitivity of error number and RT among these linguistic layers, which means that some of phonological units could have different sensitivity (or dominance) when lexical encoding in interactive manner.

Chapter 5
Conclusion

With the result of speech error and response time, we could respond to the research questions proposed in charter 2. First, by means of controlling phonological similarity, the speed of response displayed diverse patterns when we compared these tasks in pairs, even though the phonological similarity did not induce significant difference within each task. Dell's model provided a theoretical basis for dual inputs and why the results were different when subjects received different visual tasks.

Second, with regard to linguistic effects, we found that Stroop effect, syllable structure, phonotactic regularity, and tone effects were significant in the generation of speech errors. These effects seem to impact on the lexical process apparently in these tests.

Third, concerning the issue of advance planning unit, we concluded that the unit of tonal syllable and syllable could serve as possible units in lexical planning, while the units of onset, vowel, and rhyme might be not significant enough to reduce speech errors or process faster in lexical production.

Fourth, the status of tone seems to be a lexicon-like unit because it did not bring out any facilitative effect in processing speed but helped reduce the error amount of speech errors. The possible account is that tone might be a lexical organization and encoded in phonological representation, not a pure phonological tone.

Fifth, under the visual competition of color term and visual concept, the dual inputs and the diverse results among different visual tasks seem to support interactive account of lexical access, because interactive account provides a flexible and theoretical basis to explain Stroop effect in current study.

For future study, we could combine Stroop technique, and controlling of phonological similarity as well as the semantic domain to examine the lexical access in aphasic speech or in children's language development. Phonological similarity might induce different patterns of speech errors and response time in aphasic lexical network or the one which is developing language. By means of Stroop technique, we could also extend the domain outside the color, such as naming of number system to testify operation of processing model. However, we still need more psychological evidence to distinguish the processes of naming and reading task, even though the data collected so far could imply the different mechanisms between them. We need more direct evidence, such as ERP or MEG tests, to help us explain the operation between linguistic mechanism and the visual tasks.

At the end of this study, we supposed that classical experiment of psychological technique could provide linguistics more evidence and insights on the interaction among linguistic representations, as well as the lexical structure in mental process. Stroop technique is an ideal case to open the cross-field view of linguistics and psychology.

Reference

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short term memory. *Journal of Verbal Learning and Verbal Behavior*, 14, 575–589.

Berg, T. (2005). A Structural Account of Phonological *Paraphasias. Brain and Language*, 94, 104-129.

Boomer, D. S. & Laver, J. D. M. (1968). Slips of the tongue. British Journal of Disorders of Communication 3, 2-12.

Buckingham, H. W. (1980). On correlating aphasic errors with slips-of-tongue. *Applied Psycholinguistics, 1*, 199-220.

Butterworth, B. L. (1980). Evidence from pauses in speech. In B. Butterworth (Ed.), *Language production* (Vol. 1, pp. 155-176). London: Academic Press.

Butterworth, B. L. (1982). Speech errors: Old date in search of new theories. In A. Culter (Ed.), Slips of tongue and language production. Amsterdam: Mouton.

Chen, J.-Y. (1999). The representation and processing of tone in Mandarin Chinese: Evidence from slips of the tongue. Applied Psycholinguistics, 20, 289-301.

Chen, J.-Y. & Dell, G. S. (2003). Word Form Encoding in Chinese Speech Production. Chinese Journal of Psychology, 45 (4), 313-322.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review 93 (3)*, 283-321.

Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. Journal of Memory and Language 27, 124-142.

Dell, G. S.& Gordon, J. K. (2003). Neighbors in the lexicon: Friends or foes? In N. O. Schiller and A. S. Meyer (Eds.), *Phonetics and phonology in language* *comprehension and production: Differences and similarities*. New York: Mouton de Grugter.

Dell, G. S.& Juliano, C. (1996). Computational models of phonological encoding. In T. Dijkstra & K. de Smedt (Eds.) *Computational psycholinguistics: AI and connectionist models of human language processing* (pp. 328-359). Taylor & Francis.

Dell, G. S.& Juliano & Govindjee, A. (1993). Structure and content in language production: A theory of frame constraints in phonological speech errors. *Cognitive Science, 17*, 149-195.

Dell, G. S. & O'Seaghdha, P. G. (1991). Mediated and convergent lexical priming in language production: A comment on Levelt at al. (1991). *Psychological Review*.

Dell, G. S. & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. Cognition 42, 287-314.

Dell, G. S. & Reich, P. A. (1981) Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior, 20*, 611-629.

Dell, G. S., & Repka, R. J. (1992). Errors in inner speech. In B. J. Baars (Ed.), Experimental *slips and human error: Exploring the architecture of volition*, 237–262. New York: Plenum.

Dell, G. S. & Svec, W. R. (1997). Language Production and Serial Order: A Functional Analysis and a Model, *Psychological Review*, 104 (1), 123-147.

Dell, G. S. & Schwartz, M. F. & Martin, N. & Gagnon, D. A. (1997). Lexical Access in Aphasic and Nonaphasic Speakers. *Psychological Review*, 104 (4), 801-838.

Dell, G. S., Lawler, E. N., Harris, H. D., & Gordon, J. K. (2004). Models of errors of omission in aphasic naming. *Cognitive Neuropsychology*, 21, 125–145.

Fay, D. & Cutler, A. (1977). *The Phonology of Standard Chinese.* Oxford University Press.

Forster, K. I. (1976). Accessing the mental lexicon. In F. J. Wales & E. Walker (Eds.) *New Approaches to language mechanisms* (pp.257-287). Amsterdam: North-Holland.

Fromkin, V. (1971). The nonanomalous nature of anomalous utterances. *Linguistics* 4, 47-68.

Fromkin (1973). *Speech errors as linguistic evidence*. The Hauge: Mouton.

Fromkin, V. & Ratner, N. B. (1998). Speech production. In J. B. Gleason & N. B. Ratner (Eds.), Psycholinguistics (pp. 309-346). California: Wadsworth.

Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Speech production (Vol. 1*). New York: Academic Press.

Garrett, M. F. (1988). Processes in language production. In F. J. Newmeyer (Ed.), Linguistics*: The Cambridge Survey III. Language: Psychological and biological aspects* (pp.69-96). Cambridge: Cambridge University Press.

Gordon, Jean K. (2002). Phonological neighborhood effects in aphasic speech errors: spontaneous and structured contexts. *Brain and Language*, 82, 113-145.

Hockett, C. F. (1967). Where the tongue slips, there slip I. In V. Fromkin (Ed.), Speech errors as linguistic evidence (pp.93-119). Netherlands: Mouton, the Hague.

Hotopf, W. H. N. (1980). Semantic similarity as a factor in whole-word slips of the tongue. In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand*, 97–109. New York: Academic Press.

Huang, C. T. James, Li, Y.-H. Audrey & Li, Y. (2008). *The syntax of Chinese.* Cambridge University Press.

Levelt, W. J. M. (1989). *Speaking: From intension to articulation*. Cambridge: MIT Press.

Levelt, W. J. M. (ed.) (1993). *Lexical Access in Speech Production*. UK: Blackwell

Publisher.

Levelt, W. J. M. (1993). Accessing Words in Speech Production: Stages, Processes and Representations. *Lexical Access in Speech Production,* 1-22.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–75.

Levitt, A.G. & Healy, A.F. (1985). The roles of phonese frequency, similarity, and availability in the experimental elicitation of speech errors. *Journal of Memory and Language 24*, 717-733.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: the Neighborhood Activation Model. *Ear and Hearing*, 1–36.

Mackay, D. G. (1970). Spoonerisms: The structure of errors in the serial order of speech. *Neuropsychologia 8*, 323-350.

Mackay, D. G. (1972). The structure of wrds and syllables: Evidence from errors in speech. *Cognitive Psychology*, *3*, 210-227.

Maddieson, Ian (1986). The size and structural of phonological inventories: analysis of UPSID. In Ohala, John & Jeri Jaeger (eds), *Experimental Phonology*. Academic Press.

Martin, N., Gagnon, D. A., Schwartz, M. F., Dell, G. S. & Saffran, E. M. (1996). Phonological facilitation of semantic errors in normal and aphasic speakers. *Language and Cognitive Processes 11* (3), 257-282.

Martin, N., & Ayala, J. (2004). Measurements of auditoryverbal STM in aphasia: effects of task, item and word processing impairment. *Brain and Language*, *8*, 464–483.

McClelland & Rumelhart (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review 88* (5), 375-407.

Meringer, R. & Mayer, K. (1895). *Versprechen und Verlesen: eine* psychologisch-*linguistische Studie.* Stuttgart: Goschense Verlagsbuchhandlung.

Meyer, A. S. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language*, *29*, 524-545.

Meyer, A. S. (1991). The time course of phonological encoding in language production: Phonological encoding inside a syllable. *Journal of Memory and Language, 30*, 69-89.

Nooteboom, S.G. (1969). The tongue slips into patterns. *Leyden Studies in Linguistics* and *Phonetics*. The Hague: Mouton.

Oppenheim, G. M. & Dell, G. S. (2008). Inner speech slips exhibit lexical bias, but not the phonemic similarity effect. *Cognition*, 106, 528-537.

Packard, J. L. (1986). Tone Production deficits in nonfluent aphasic Chinese speech, *Brain and Language*, 29, 212-223.

Rapp, D. N. & Samuel, A.G. (2002). A reason to rhyme: Phonological and semantic influences on lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition 28* (3), 564-571.

Roelofs, A. (1996). Computationa; models of lemma retrieval. In T. Dijkstra & K. de Smedt (Eds.) *Computational Psycholinguistics: AI and connectionist mode4ls of human language processing* (pp.308-327). Taylor & Francis.

Roelofs, A. (1997) The WEAVER model of word-form encoding in speech production.

Roelofs, A., & Meyer, A.S. (1998). Metrical structure in planning the production of spoken words. Journal of Experimental Psychology: *Learning, Memory, and Cognition*, *24*, 922–939.

Schwartz, M. F. & Dell, G. S. & Martin, N. (2004). Testing the interactive two-step model of lexical access: Part I. picture naming, *Brain and Language*, 91, 71-72.

Sevald, C. A., Dell, G. S., &Cole, J. (1995). Syllable structure in speech production: Are syllable chunks or schemas? *Journal of Memory and Language, 34*, 807-820.

Shattuck-Hufnagel (1979). Speech errors as evidence for a serial ordering mechanism in sentence production. In W. E. Cooper & E. C. T. Walkers (Eds.), *Sentence processing* (pp.295-342). Hillsdale, N. J.: Lawrence Erlbaum Assiciates.

Shattuck-Hufnagel (1986). The representation of phonologoical information during speech production planning: Evidence from vowel errors in spontaneous speech. *Phonology Yearbook 3*, 117-149.

Shattuck-Hufnagel (1987). The role of word onset consonants in speech production planning: New evidence from speech error patterns. In E. Keller & M. Gopnik (Eds.), *Motor and sensory processing in language* (pp. 17-51). Hillsdale, NJ: Erlbaum.

Schwartz, M. F. (1994). Disordered Speech Production in Aphasic and Normal Speakers. *Brain and Language*, 47, 52-88.

Sokolov, A. N. (1972). Inner speech and thought. New York: Plenum.

Stemberger, J. P. (1983). Speech errors and theoretical phonology: A review. Bloomington: Indiana Linguistics Club.

Stemberger, J. P. (1984). Structural Errors in Normal and Agrammatic Speech. *Cognitve Neuropsychology*, 1 (4), 281-313.

Stemberger, J. P. (1985). An interactive activation model of language production. In W. E. Andrew (Ed.), Progress in the psychology of language (pp. 143-186). London: Lawrence Erlbaum Associates Ltd., Publishers.

Stemberger, J. P. (1989). Speech errors in early child language production. *Journal of Memory and Language, 28(2)*, 164-188.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. Journal of Experimental Psychology, 18, 643-662.

Vitevitch, M. S. (1997). The neighborhood characteristics of malapropisms. *Language and Speech*, 40, 211–228.

Vitevitch, M. S., & Luce, P. A. (1998). When words compete: levels of processing in perception of spoken words. *Psychological Science*, 9, 325–329.

Vitevitch, M. S. (2002, July). The influence of phonological similarity neighborhoods on speech production. *J Exp Psychol Learn Mem Cogn*, 28(4), 735–747.

Vitevitch, M. S., Armbruster, J., & Chu, S. (2004). Sublexical and lexical representations in speech production: effects of phonotactic probability and onset density. *J Exp Psychol Learn Mem Cogn*, 30(2), 514–529.

Vitevitch, M. S., & Sommers, M. S. (2003). The facilitative influence o f phonological similarity and neighborhood frequency in speech production in younger and older adults. *Memory and Cognition*, 31, 491–504.

Vitevitch, M. S., & Stamer, M. K. (2006). The curious case of competition in Spanish speech production. *Language & Cognitive Processes*, 21(6), 760–770.

Wan, I. P. & Jeri J. (1998) Speexh errors and the representation of tone in Mandarin Chinese. *Phonology*, 15, 417-461.

Wan, I. P. (1999). Mandarin phonology: Evidence from speech errors. Ph D. Dissertation.

Wan, I. P. (2007). Aphonological Investigation in Speech Errors and Aphasic Speech in Mandarin. Taipei: Crane.

Wells, R. (1951). Predicting slips of the tongue. *Yale Scientific Magazine*, *3*, 9-30.

Appendix 1: Stimuli Design

Trials design in Task 1

Color Naming Test					
H1	黑灰 黃紅 灰紅 黃灰		L1	綠白 黃紫 綠黃 橙綠	
H2	黃紅 灰棕 紅灰 黃棕		L2	紫白 灰紫 白綠 橙灰	
H3	棕紅 灰黑 紅黃 灰黃		L3	黑藍 橙紫 綠黑 黃藍	
H4	藍白 黑灰 黑紅 棕黃		L4	紫橙 黃綠 黑藍 橙黑	
H5	黃紅 灰黃 藍白 黑紅		L5	黃藍 紫橙 綠黃 橙綠	
H6	紅黃 黃棕 灰黑 紅灰		L6	綠黃 橙綠 白紫 黃綠	
H7	紅灰 灰棕 黃紅 棕灰		L7	紫藍 黃橙 綠紫 藍黑	
H8	黑紅 黃灰 紅灰 灰棕		L8	黑橙 綠黃 藍紫 黑藍	
H9	棕黃 黑灰 紅棕 白藍		L9	紫黃 黑橙 綠橙 藍黑	
H10	灰黃 灰棕 紅灰 黃黑		L10	黃紫 綠橙 藍橙 黃綠	
H11	灰黃 紅灰 紅黃 灰黑		L11	綠橙 黃綠 紫黃 白綠	
H12	棕黃 灰紅 棕灰 紅黃		L12	灰橙 綠白 紫灰 白紫	
H13	黃灰 黃紅 黑灰 紅棕		L13	藍黃 黑綠 紫橙 藍黑	
H14	黃棕 紅黑 灰黑 白藍		L14	黑橙 藍黑 綠黃 橙紫	
H15	紅黑 白藍 黃灰 紅黃		L15	橙紫 綠橙 黃綠 藍黃	
H16	灰紅 黑灰 棕黃 黃紅		L16	綠黃 紫白 綠橙 黃綠	
H17	灰棕 紅黃 棕灰 灰紅		L17	黑藍 紫綠 橙黃 藍紫	
H18	棕灰 灰紅 灰黃 紅黑		L18	藍黑 紫藍 黃綠 橙黑	
H19	藍白 棕紅 灰黑 黃棕		L19	黑藍 橙綠 橙黑 黃紫	
H20	紅黃 灰紅 棕灰 黃灰		L20	綠黃 橙藍 橙綠 紫黃	

Trials design in Task 2

Color Reading Test			
H1	黑灰 黃紅 灰紅 黃灰	L1	綠白 黃紫 綠黃 橙綠
H2	黃紅 灰棕 紅灰 黃棕	L2	紫白 灰紫 白綠 橙灰
H3	棕紅 灰黑 紅黃 灰黃	L3	黑藍 橙紫 綠黑 黃藍
H4	藍白 黑灰 黑紅 棕黃	L4	紫橙 黃綠 黑藍 橙黑
H5	黃紅 灰黃 藍白 黑紅	L5	黃藍 紫橙 綠黃 橙綠
H6	紅黃 黃棕 灰黑 紅灰	L6	綠黃 橙綠 白紫 黃綠
H7	紅灰 灰棕 黃紅 棕灰	L7	紫藍 黃橙 綠紫 藍黑
H8	黑紅 黃灰 紅灰 灰棕	L8	黑橙 綠黃 藍紫 黑藍
H9	棕黃 黑灰 紅棕 白藍	L9	紫黃 黑橙 綠橙 藍黑
H10	灰黃 灰棕 紅灰 黃紅	L10	黃紫 綠橙 藍橙 黃綠
H11	灰黃 紅灰 紅黃 灰黑	L11	綠橙 黃綠 紫黃 白綠
H12	棕黃 灰紅 棕灰 紅黃	L12	灰橙 綠白 紫灰 白紫
H13	黃灰 黃紅 黑灰 紅棕	L13	藍黃 黑綠 紫橙 藍黑
H14	黃棕 紅黑 灰黑 白藍	L14	黑橙 藍黑 綠黃 橙紫
H15	紅黑 白藍 黃灰 紅黃	L15	橙紫 綠橙 黃綠 藍黃
H16	灰紅 黑灰 棕黃 黃紅	L16	綠黃 紫白 綠橙 黃綠
H17	灰棕 紅黃 棕灰 灰紅	L17	黑藍 紫綠 橙黃 藍紫
H18	棕灰 灰紅 灰黃 紅黑	L18	藍黑 紫藍 黃綠 橙黑
H19	藍白 棕紅 灰黑 黃棕	L19	黑藍 橙綠 橙黑 黃紫
H20	紅黃 灰紅 棕灰 黃灰	L20	綠黃 橙藍 橙綠 紫黃

Trials design in Task 3

Stroop Naming Test			
H1	黑灰 黃紅 灰紅 黃灰	L1	綠白 黃紫 綠黃 橙綠
H2	黃紅 灰棕 紅灰 黃棕	L2	紫白 灰紫 白綠 橙灰
H3	棕紅 灰黑 紅黃 灰黃	L3	黑藍 橙紫 綠黑 黃藍
H4	藍白 黑灰 黑紅 棕黃	L4	紫橙 黃綠 黑藍 橙黑
H5	黃紅 灰黃 藍白 黑紅	L5	黃藍 紫橙 綠黃 橙綠
H6	紅黃 黃棕 灰黑 紅灰	L6	綠黃 橙綠 白紫 黃綠
H7	紅灰 灰棕 黃紅 棕灰	L7	紫藍 黃橙 綠紫 藍黑
H8	黑紅 黃灰 紅灰 灰棕	L8	黑橙 綠黃 藍紫 黑藍
H9	棕黃 黑灰 紅棕 白藍	L9	紫黃 黑橙 綠橙 藍黑
H10	灰黃 灰棕 紅灰 黃紅	L10	黃紫 綠橙 藍橙 黃綠
H11	灰黃 紅灰 紅黃 灰黑	L11	綠橙 黃綠 紫黃 白綠
H12	棕黃 灰紅 棕灰 紅黃	L12	灰橙 綠白 紫灰 白紫
H13	黃灰 黃紅 黑灰 紅棕	L13	藍黃 黑綠 紫橙 藍黑
H14	黃棕 紅黑 灰黑 白藍	L14	黑橙 藍黑 綠黃 橙紫
H15	紅黑 白藍 黃灰 紅黃	L15	橙紫 綠橙 黃綠 藍黃
H16	灰紅 黑灰 棕黃 黃紅	L16	綠黃 紫白 綠橙 黃綠
H17	灰棕 紅黃 棕灰 灰紅	L17	黑藍 紫綠 橙黃 藍紫
H18	棕灰 灰紅 灰黃 紅黑	L18	藍黑 紫藍 黃綠 橙黑
H19	藍白 棕紅 灰黑 黃棕	L19	黑藍 橙綠 橙黑 黃紫
H20	紅黃 灰紅 棕灰 黃灰	L20	綠黃 橙藍 橙綠 紫黃

Trials design in Task 4

Homophonous Stroop Naming Test			
H1	嘿輝 皇洪 輝洪 皇輝	L1	慮白 皇籽 慮皇 徽慮
H2	皇洪 輝鬃 洪輝 皇鬃	L2	籽白 輝籽 白慮 徽輝
H3	鬃洪 輝嘿 洪皇 輝皇	L3	嘿蘭 徽籽 慮嘿 皇蘭
H4	蘭白 嘿輝 嘿洪 鬃皇	L4	籽徽 皇慮 嘿蘭 徽嘿
H5	皇洪 輝皇 蘭白 嘿洪	L5	皇蘭 籽徽 慮皇 徽慮
H6	洪皇 皇鬃 輝嘿 洪輝	L6	慮皇 徽慮 白籽 皇慮
H7	洪輝 輝鬃 皇洪 鬃輝	L7	籽蘭 皇徽 慮籽 蘭嘿
H8	嘿洪 皇輝 洪輝 輝鬃	L8	嘿徽 慮皇 蘭籽 嘿蘭
H9	鬃皇 嘿輝 洪鬃 白蘭	L9	籽皇 嘿徽 慮徽 蘭嘿
H10	輝皇 輝鬃 洪輝 皇洪	L10	皇籽 慮徽 蘭徽 皇慮
H11	輝皇 洪輝 洪皇 輝嘿	L11	慮徽 皇慮 籽皇 白慮
H12	鬃皇 輝洪 鬃輝 洪皇	L12	輝徽 慮白 籽輝 白籽
H13	皇輝 皇洪 嘿輝 洪鬃	L13	蘭皇 嘿慮 籽徽 蘭嘿
H14	皇鬃 洪嘿 輝嘿 白蘭	L14	嘿徽 蘭嘿 慮皇 徽籽
H15	洪嘿 白蘭 皇輝 洪皇	L15	徽籽 慮徽 皇慮 蘭皇
H16	輝洪 嘿輝 鬃皇 皇洪	L16	慮皇 籽白 慮徽 皇慮
H17	輝鬃 洪皇 鬃輝 輝洪	L17	嘿蘭 籽慮 徽皇 蘭籽
H18	鬃輝 輝洪 輝皇 洪嘿	L18	蘭嘿 籽蘭 皇慮 徽嘿
H19	蘭白 鬃洪 輝嘿 皇鬃	L19	嘿蘭 徽慮 徽嘿 皇籽
H20	洪皇 輝洪 鬃輝 皇輝	L20	慮皇 徽蘭 徽慮 籽皇

Trials design in Task 5

Shared Unit	**Shared Unit Test**			
Onset	H1	壺禍 毀忽 禍忽 毀禍	L1	爐補 毀卒 爐毀 侈爐
	H2	毀忽 禍宰 忽禍 毀宰	L2	卒補 禍卒 補爐 侈禍
	H3	宰忽 禍壺 忽毀 禍毀	L3	壺撈 侈卒 爐壺 毀撈
	H4	撈補 壺禍 壺忽 宰毀	L4	卒侈 毀爐 壺撈 侈壺
	H5	毀忽 禍毀 撈補 壺忽	L5	毀撈 卒侈 爐毀 侈爐
	H6	忽毀 毀宰 禍壺 忽禍	L6	爐毀 侈爐 補卒 毀爐
	H7	忽禍 禍宰 毀忽 宰禍	L7	卒撈 毀侈 爐卒 撈壺
	H8	壺忽 毀禍 忽禍 禍宰	L8	壺侈 爐毀 撈卒 壺撈
	H9	宰毀 壺禍 忽宰 補撈	L9	卒毀 壺侈 爐侈 撈壺
	H10	禍毀 禍宰 忽禍 毀忽	L10	毀卒 爐侈 撈侈 毀爐
Vowel	H1	黑灰 告狗 灰狗 告灰	L1	徐反 告釋 徐告 奮徐
	H2	告狗 灰割 狗灰 告割	L2	釋反 灰釋 反徐 奮灰
	H3	割狗 灰黑 狗告 灰告	L3	黑擦 奮釋 徐黑 告擦
	H4	擦反 黑灰 黑狗 割告	L4	釋奮 告徐 黑擦 奮黑
	H5	告狗 灰告 擦反 黑狗	L5	告擦 釋奮 徐告 奮徐
	H6	狗告 告割 灰黑 狗灰	L6	徐告 奮徐 反釋 告徐
	H7	狗灰 灰割 告狗 割灰	L7	釋擦 告奮 徐釋 擦黑
	H8	黑狗 告灰 狗灰 灰割	L8	黑奮 徐告 擦釋 黑擦
	H9	割告 黑灰 狗割 反擦	L9	釋告 黑奮 徐奮 擦黑
	H10	灰告 灰割 狗灰 告狗	L10	告釋 徐奮 擦奮 告徐
Syllable	H1	黑迴 慌哄 迴哄 慌迴	L1	驢敗 慌滋 驢慌 逞驢
	H2	慌哄 迴粽 哄迴 慌粽	L2	滋敗 迴滋 敗驢 逞迴
	H3	粽哄 迴黑 哄慌 迴慌	L3	黑覽 逞滋 驢黑 慌覽
	H4	覽敗 黑迴 黑哄 粽慌	L4	滋逞 慌驢 黑覽 逞黑
	H5	慌哄 迴慌 覽敗 黑哄	L5	慌覽 滋逞 驢慌 逞驢
	H6	哄慌 慌粽 迴黑 哄迴	L6	驢慌 逞驢 敗滋 慌驢
	H7	哄迴 迴粽 慌哄 粽迴	L7	滋覽 慌逞 驢滋 覽黑
	H8	黑哄 慌迴 哄迴 迴粽	L8	黑逞 驢慌 覽滋 黑覽
	H9	粽慌 黑迴 哄粽 敗覽	L9	滋慌 黑逞 驢逞 覽黑
	H10	迴慌 迴粽 哄迴 慌哄	L10	慌滋 驢逞 覽逞 慌驢
Rhyme	H1	累備 趙冬 備冬 趙備	L1	距海 趙癡 距趙 亨距
	H2	趙冬 備拱 冬備 趙拱	L2	癡海 備癡 海距 亨備
	H3	拱冬 備累 冬趙 備趙	L3	累閃 亨癡 距累 趙閃
	H4	閃海 累備 累冬 拱趙	L4	癡亨 趙距 累閃 亨累
	H5	趙冬 備趙 閃海 累冬	L5	趙閃 癡亨 距趙 亨距
	H6	冬趙 趙拱 備累 冬備	L6	距趙 亨距 海癡 趙距
	H7	冬備 備拱 趙冬 拱備	L7	癡閃 趙亨 距癡 閃累

	H8	累冬 趟犕 冬犕 犕拱	L8	累亨 距趟 閃癡 累閃
	H9	拱趟 累犕 冬拱 海閃	L9	癡趟 累亨 距亨 閃累
	H10	犕趟 犕拱 冬犕 趟冬	L10	趟癡 距亨 閃亨 趟距
Tone	H1	屋貪 凡離 貪離 凡貪	L1	敗初 凡普 敗凡 綢敗
	H2	凡離 貪苛 離貪 凡苛	L2	普初 貪普 初敗 綢貪
	H3	苛離 貪屋 離凡 貪凡	L3	屋足 綢普 敗屋 凡足
	H4	足初 屋貪 屋離 苛凡	L4	普綢 凡敗 屋足 綢屋
	H5	凡離 貪凡 足初 屋離	L5	凡足 普綢 敗凡 綢敗
	H6	離凡 凡苛 貪屋 離貪	L6	敗凡 綢敗 初普 凡敗
	H7	離貪 貪苛 凡離 苛貪	L7	普足 凡綢 敗普 足屋
	H8	屋離 凡貪 離貪 貪苛	L8	屋綢 敗凡 足普 屋足
	H9	苛凡 屋貪 離苛 初足	L9	普凡 屋綢 敗綢 足屋
	H10	貪凡 貪苛 離貪 凡離	L10	凡普 敗綢 足綢 凡敗
Tonal syllable	H1	黑燁 皇洪 燁洪 皇燁	L1	慮百 皇籽 慮皇 懲慮
	H2	皇洪 燁鬃 洪燁 皇鬃	L2	籽百 燁籽 百慮 懲燁
	H3	鬃洪 燁黑 洪皇 燁皇	L3	嘌蘭 懲籽 慮嘌 皇蘭
	H4	蘭百 嘌燁 嘌洪 鬃皇	L4	籽懲 皇慮 嘌蘭 懲嘌
	H5	皇洪 燁皇 蘭百 嘌洪	L5	皇蘭 籽懲 慮皇 懲慮
	H6	洪皇 皇鬃 燁嘌 洪燁	L6	慮皇 懲慮 百籽 皇慮
	H7	洪燁 燁鬃 皇洪 鬃燁	L7	籽蘭 皇懲 慮籽 蘭嘌
	H8	嘌洪 皇燁 洪燁 燁鬃	L8	嘌懲 慮皇 蘭籽 嘌蘭
	H9	鬃皇 嘌燁 洪鬃 百蘭	L9	籽皇 嘌懲 慮懲 蘭嘌
	H10	燁皇 燁鬃 洪燁 皇洪	L10	皇籽 慮懲 蘭懲 皇慮

Appendix 2: Overview of Results

Task 1	RT		Trial F		Error N	
	High	Low	High	Low	High	Low
S1	4464.30	4623.55	0	1	0	1
S2	5405.20	5568.00	0	2	0	2
S3	4767.20	4853.35	1	1	1	0
S4	5593.25	5404.60	2	0	2	0
S5	5935.05	5705.50	2	0	1	1
S6	7271.35	7106.90	2	2	2	2
S7	5019.90	5630.50	1	1	1	1
S8	4703.60	5078.90	1	5	1	3
S9	5598.50	5591.75	2	0	1	1
S10	5117.40	4861.70	0	0	0	0
S11	5900.80	6102.75	3	1	2	1
S12	5202.55	4851.30	6	1	6	1
S13	6730.60	7383.30	3	5	0	5
S14	4812.80	4763.00	1	1	1	1
S15	4930.35	5184.35	3	4	3	4
S16	4262.75	3892.15	4	2	5	0
S17	4331.20	4972.45	1	4	0	4
S18	5273.00	5687.05	1	3	0	3
S19	3765.10	3617.25	3	0	2	1
S20	8109.37	8511.30	7	6	7	3
S21	5889.45	5173.65	5	1	1	4
S22	6295.50	6628.05	4	4	1	5
Average	5384.94	5619.85	52.00	44.00	37.00	43.00

Task 2	RT		Trial F		Error N	
	High	Low	High	Low	High	Low
S1	2940.55	2714.05	1	2	0	1
S2	4039.20	3864.45	0	3	0	2
S3	3826.20	3345.75	4	1	4	1
S4	3653.50	3700.45	0	0	0	0
S5	3144.68	2952.05	1	0	1	0
S6	4791.10	4516.60	2	1	2	1
S7	4483.95	4486.60	3	1	3	1
S8	3157.85	3038.10	7	0	7	0
S9	3111.45	3025.80	1	0	1	0
S10	4652.35	4604.35	0	0	0	0
S11	3403.05	3239.90	1	0	0	1
S12	2898.30	2907.75	1	3	1	3
S13	4019.80	3714.60	2	0	2	0
S14	3125.70	2836.00	4	0	2	0
S15	4331.50	4408.30	1	0	1	0
S16	2737.05	2496.10	6	1	5	0
S17	3729.00	3404.60	5	0	4	0
S18	3933.80	3979.10	0	0	0	0
S19	2452.40	2090.35	3	1	2	0
S20	5343.45	5244.65	7	8	7	7
S21	3294.05	2889.00	3	2	3	2
S22	4866.60	4840.20	0	3	0	2
Average	3724.34	3559.03	52	26	45.00	21.00

Task 3	RT		Trial F		Error N	
	High	Low	High	Low	High	Low
S1	6110.20	5546.35	6	0	6	0
S2	6703.55	6421.00	8	0	8	0
S3	4756.65	4723.20	2	2	2	2
S4	6260.85	6778.85	0	2	0	2
S5	6700.70	6986.45	4	2	2	2
S6	7306.75	8552.75	2	13	2	13
S7	8151.35	8062.45	13	5	14	4
S8	8506.70	8622.20	5	12	5	12
S9	6035.35	6322.60	5	1	6	0
S10	5985.45	6174.15	3	5	3	5
S11	6738.60	7329.00	2	0	2	0
S12	6405.30	6598.90	3	2	3	2
S13	7240.05	6871.60	8	6	8	5
S14	4801.60	5488.52	1	5	1	5
S15	5386.70	5235.15	4	3	4	3
S16	5603.40	5690.05	5	3	5	1
S17	6259.55	7043.80	6	6	6	6
S18	7005.10	6439.55	11	0	11	0
S19	5050.21	4980.35	7	4	7	4
S20	7728.75	7723.05	10	9	9	9
S21	7284.40	6956.90	27	12	27	10
S22	8409.05	8284.90	20	13	19	12
Average	6565.01	6674.17	152	105	150.00	97.00

Task 4	RT		Trial F		Error N	
	High	Low	High	Low	High	Low
S1	4740.35	4829.15	0	0	1	1
S2	5304.70	5134.15	1	6	1	6
S3	4584.65	4592.55	2	3	2	3
S4	6768.70	7395.95	4	0	4	0
S5	7142.45	8146.10	10	7	9	6
S6	6689.15	6918.30	5	13	5	12
S7	6946.95	7612.50	1	5	1	5
S8	6734.35	7124.95	5	8	5	7
S9	5651.90	6054.48	7	5	7	4
S10	6147.55	6551.65	2	6	2	6
S11	7239.60	7231.15	1	3	1	2
S12	6674.50	6262.25	15	7	17	5
S13	7760.15	8110.50	4	6	3	4
S14	4592.85	4527.20	1	0	1	0
S15	5293.10	5240.70	10	7	11	5
S16	4730.80	4545.10	5	4	6	1
S17	6154.50	7151.20	5	8	6	7
S18	6328.25	6597.25	2	5	2	4
S19	4839.85	4877.45	0	4	0	4
S20	6784.40	7336.00	13	12	13	11
S21	6302.20	6101.80	12	7	8	9
S22	8633.65	8879.25	13	15	11	14
Average	6183.85	6419.07	118	131	116.00	116.00

Task 5	Onset		Vowel		Rhyme	
	RT	Error N	RT	Error N	RT	Error N
S1	5432.15	3	5512.55	3	5017.90	0
S2	4778.25	2	4603.20	2	4580.90	2
S3	4509.40	3	4189.10	0	4234.25	0
S4	6409.20	0	7296.26	0	6194.60	2
S5	6679.00	3	5506.05	1	5979.25	2
S6	6281.15	5	6858.70	8	6417.65	7
S7	5365.55	4	5985.35	4	5780.55	2
S8	5706.05	1	6033.60	5	5018.20	2
S9	5369.70	3	5084.80	3	4846.30	6
S10	5432.90	1	5385.35	0	5391.50	1
S11	5525.11	1	5591.65	0	5206.60	2
S12	5813.60	2	5263.70	3	5116.20	3
S13	6263.25	3	6640.65	7	6215.70	4
S14	4077.65	0	4419.95	3	4523.65	3
S15	4762.90	0	4750.50	2	4723.70	2
S16	4277.85	4	4128.05	3	4470.35	4
S17	5306.95	1	5105.15	3	5220.60	3
S18	6117.50	3	5503.65	2	5054.85	4
S19	3582.75	1	4126.90	1	4033.40	2
S20	7235.05	12	6726.55	5	7270.55	8
Total	5446.30	52.00	5435.59	55.00	5264.84	59.00

Task 5	Tone		Bare Syllable		Tonal Syllable	
	RT	Error N	RT	Error N	RT	Error N
S1	5363.85	2	4435.50	5	3958.30	4
S2	4829.80	1	4605.20	1	4526.25	1
S3	4725.95	4	4043.85	2	3868.50	2
S4	6053.55	1	7287.20	1	6918.05	0
S5	5914.95	0	5252.50	3	4912.75	0
S6	7571.70	8	6369.65	6	5862.40	6
S7	6051.75	6	5048.70	0	4947.40	0
S8	5478.90	0	5210.55	1	4406.15	3
S9	5248.80	4	4418.50	4	3683.45	3
S10	5376.55	0	5169.85	2	4918.00	0
S11	5858.75	2	4574.70	0	4283.70	0
S12	4821.00	1	4606.70	2	4428.85	1

S13	6268.10	3	5918.10	6	5273.25	0
S14	4273.05	1	4427.25	4	3922.50	2
S15	4938.55	2	4662.75	0	4408.10	0
S16	4470.25	2	3690.80	5	3403.45	2
S17	5378.00	1	4967.15	1	4269.45	1
S18	5409.55	2	5160.35	0	4577.30	0
S19	4066.00	3	3416.25	4	3025.00	1
S20	7304.85	5	6146.05	4	5828.65	3
Total	5470.20	0.00	4970.58	0.00	4571.08	0.00