Chapter 2
Literature Review

This chapter includes the following issues. First, the reason why slips of the tongue are taken as the research tool in the study will be introduced in 2.1. Next in 2.2, it will be reviewed how speech errors are used to construct speech production models in Germanic languages, especially focusing on the comparison with serial-ordering and connectionist model. After that, in order to investigate language processing in Mandarin thoroughly, some basic semantic and phonological properties of the language will be included in 2.3. And the relative studies will be covered in 2.4. Finally, the details of research questions will be provided in 2.5.

2.1 Slips of the tongue

The importance of slips of the tongue has been noticed since Meringer and Meyer (1895). Following their study, speech errors were collected on a large scale to investigate certain phonological theories (e.g., Gandour, 1977; Shattuck-Hufnagel, 1986; Wan & Jaeger, 2003) or to construct cognitive models (e.g., Fromkin, 1971; Garrette, 1980a; Dell & Reich, 1981; Stemberger, 1985b). Thus the nature and some specialized terms of speech errors should be reviewed first. In the following, how they are applied in previous studies will be demonstrated, including its reliability, advantages and limitations.

2.1.1 Definition

A slip of the tongue is an ‘involuntary deviation’ in a speaker’s performance from his current phonological, grammatical or lexical intention (Boomer & Laver, 1973). Such phenomenon occurs very often in our daily conversation. They are more than a
funny phenomenon but provide indirect evidence for the units, cognitive process and psychological validity of language production.

Some terms are used to define speech errors. The erroneous utterance is called 'target,' the unit which is replaced is called 'source' and the original word the speaker planned to say is labeled 'intended utterance.' The target and the replaced unit are called 'interacting units.'

Speech errors can be primarily classified into two types—phonological and lexical. Both types encompass different sizes of units. Phonological errors are those involving sounds as erroneous units, such as phonetic features, phonemic segments, consonant clusters, rhymes, tones, etc. Lexical errors are composed of morphemes, words and phrases. If the target and source are equal in the size of unit, such as one segment replaces another, it is called substitution. However, if the target gets an extra unit, it becomes an addition. Similarly, deletions mean the target unit is deleted.

Moreover, speech errors may be contextual or non-contextual. Contextual errors, also called syntagmatic errors, have a distinct source in the preceding or following context. If the source precedes the target, it is called perseveration. On the contrary, when the source shows up in the future utterance, it is marked as anticipation. And if the interacting units are purely misordered, it is called exchange. Non-contextual errors, also named paradigmatic errors, have no explicit source in the current utterance. Examples of lexical syntagmatic and paradigmatic errors are given below.

(1)(a) I wouldn’t buy macadamia nuts for the kids →

I wouldn’t kids for the macadamia nuts

(b) It’s at the bottom…I mean…top of the stack of books (from Fromkin, 1971)

(1a) is a syntagmatic lexical error, in which kids and macadamia nuts in the context exchange with each other. Conversely, (1b) is a paradigmatic lexical error since the speaker uses a semantic-related lexicon bottom, which is never shown in context, to
replace top. By the way, we can tell top is the speaker’s intended utterance because the rest of sentence after I mean... shows his intention to correct his erroneous usage of bottom.

2.1.2 Application

The reliability of speech errors lies in its wide application as evidence of linguistic theory. First, the analysis of speech errors can be used to examine the representation of phonological information in speech production (Gandour, 1977; Shattuck-Hufnagel, 1986; Wan & Jaeger, 2003; Wan, 2007a). Secondly, it can be used to build speech production models (Mackay, 1970; Fromkin, 1971; Shattuck-Hufnagel, 1979; Dell & Reich, 1981; Stemberger, 1985b; Dell, 1986; Garrett, 1988). Thirdly, the way of collecting normal speech errors have been adopted across many languages (German: Meringer & Mayer, 1895; Dutch: Cohen, 1966; Nooteboom, 1969; Spanish: Berg, 1991; Thai: Gandour, 1977; Mandarin Chinese: Chen, 1993, 1999; Wan & Jaeger, 1998, 2003; Wan, 2007a, 2007b; English: Boomer & Laver, 1968; Fromkin, 1971; Fay & Cutler, 1977; Shattuck-Hufnagel, 1986, among others) In addition, the corpus of children’s slips of the tongue are also built to investigate language development (Aitchison & Straf, 1982; Stemberger, 1989). Undoubtedly, such the wide application increases the reliability of speech error data and its further analysis.

However, people may doubt the way of collecting speech errors have some bias. For example, the accuracy of sifting speech errors may be distracted by focusing on the content simultaneously (Johnson, 1980). In fact such selective attention can be avoided through high-fidelity recordings and the multiply-checked transcription, which will be included in our ways of collecting data (see 3.1 for details).

Moreover, there is a further detectability problem that some kinds of errors are harder to be heard than others, and even the careful transcription may miss some slips.
As for this bias, Cutler (1982) submitted three arguments to verify the reliability of speech error data, accordingly ‘Some Errors,’ ‘More Errors’ and ‘No Errors’ argument. ‘Some Errors’ simply interprets the characteristics of errors. For example, movement errors suggest morphological accommodation to their new environment (Garrett, 1980a). ‘More Errors’ suggests the relative frequency of particular types of errors, such as anticipations outnumbered perseverations (Cohen, 1966; Nooteboom, 1969). ‘No Errors’ concerns errors which don’t occur, such as errors disallowed by the phonotactics of the language are almost absent (Boomer & Laver, 1968). Among these three arguments, ‘Some Errors’ argument is adopted to defense the way in comparing the relative frequency between each type of speech errors. If the differential detectability could predict the same result as the distribution of reported errors, the interpretation of data could get support (Cutler, 1982).

Nevertheless, Stemberger (1985a) mentioned some drawbacks of such observational method. First of all, the relevant error type may not be frequent enough to yield reliable results and thus the collection procedure will require a long period of time to be accomplished. Secondly, the chance estimate of each error type is not available. It will influence the accuracy of frequency counts, which is the basis of error estimates. Stemberger also noted other drawbacks such as the perceptual bias and the danger of misclassification, which will get improvement in the method (see 3.2 for details). Most importantly, he suggested that observational method is best to combine with experimental designs. As for this point, the focus of the study is to make a thorough analysis of natural corpus and the unfinished parts are left for future study.

Therefore, on the basis of the previous discussions, this study will take slips of the tongue to investigate the process of word production in Mandarin Chinese, especially the single-syllable paradigmatic lexical substitutions. These errors are important in
distinguishing two major types of speech production models since they received different explanations from each sect. The models will be illustrated more clearly in the following section.

2.2 Speech Production Models

Language production models are used to account for the major stages in language processing. They have primarily two types—serial-ordering and connectionist models. The serial-ordering model has a uni-directional processing nature while the connectionist model allows the bidirectional processing between levels. The contrasting routes of activation result in their different explanations toward paradigmatic lexical errors. In this section two types of models will be briefly introduced and compared, including their interpretation of paradigmatic errors.

2.2.1 Serial-ordering Model

The serial-ordering model, also known as discrete model, argues that the processing of language follows a series of stages step-by-step. The process will not proceed from one stage to the next unless the previous one is accomplished. In the following section, models of Fromkin (1971), Garrett (1988), and Levelt et al. (1999) will be reviewed.
2.2.1.1 Fromkin’s Utterance Generator Model

Fromkin (1971) presented the Utterance Generator model, which is able to illustrate most types of speech errors in her corpus. As Figure 2-1 has shown, it is a serial-processing model, with the top-down direction without any feedback loops.

![Figure 2-1. Fromkin’s Utterance Generator (from Fromkin, 1971)](image)

At stage 1, the meaning to be conveyed is generated. At stage 2 the syntactic structure is determined and later the semantic features are mapped accordingly. At stage 3 the sentential and phrasal stress are assigned, prior to the lexical selection at stage 4. This is evident that in many word exchanges the stress contour of the phrase is fixed by the
structure. At stage 4 the stems represented by phonemic segments are chosen, although sometimes the semantically or phonologically similar items are wrongly selected instead. Consequently their pronunciations are specified at stage 5. Finally the word is generated through motor commands.

As a pioneering model, it has several important findings. First, the model asserts segments as the minimal processing unit. It explains why segmental errors outnumber feature errors. Secondly, it also accounts for many features of speech errors, such as the phrasal stress is not destroyed by word exchanges.

However, there are still some points not clearly verified in this model. For example, the model does not illustrate the fact that major categories (nouns, verbs, adjectives) are not moved or shifted while minor categories (inflection morphemes) are (Fromkin & Ratner, 1998). Moreover, it mentions word substitutions tend to share phonological or semantic features because the wrong item is selected at stage 4, but the procedure for lexical selection is not clearly specified in the model.

2.2.1.2 Garrett’s Two-step Retrieval Process

Also based on speech errors, Garrett (1988) suggested a two-step lexical retrieval process, which is represented in Figure 2-2. It illustrates that the F (functional) and P (positional) levels are independent from each other, i.e. the retrieval of meaning and form are separated. It means at the functional level the grammatical category is assigned and the lexicon is retrieved. This is where word exchanges with the same grammatical category occur. At the positional level, sounds are assigned to lexicons and sound exchanges, stranding exchanges and morpheme shifts may occur. Therefore, the information about word form is not available at lexical selection stage, and the form of a retrieved lexicon will not be specified until certain point at the positional level.
Based on the view of distinct processing levels, Garrett (1988) mentioned several points to interpret paradigmatic lexical errors. First, he showed that the overall featural similarity of phonologically-related errors were higher than that of semantic errors and thus concluded that semantic errors were not contaminated by sounds. However, I think the scope of data should be enlarged to all paradigmatic errors rather than the comparison between errors related in sound and meaning.

*Figure 2-2.* Garrett’s two-step lexical retrieval process (from Garrett, 1988)
Secondly, he cited many experimental studies to argue that the phonological effects on semantic errors were not robust. For example, he found no consistent phonological similarity in errors of certain semantic field such as body parts and clothing terms (Garrett, 1987). However, I think such experiment has some methodological bias. For instance, the items within one particular semantic field, such as body parts, may be too limited to have evenly phonetic distribution. With such natural bias, it is no wonder that the phonological effect within semantic substitutions is not significant.

Thirdly, for mixed errors like read/write or lobster/oyster, he approved the interaction as an effect of conceptual facilitation on form selection, rather than phonological effect intruding in lexical selection. In his opinion, at the functional level the semantic node was selected. If any of them happened to be similar in form, they probably will be chosen as the output at the positional level. However, this proposal only predicts a small number of mixed errors. When the proportion of mixed errors precedes the chance level, the explanation will seem to be relatively weaker.

2.2.1.3 Levelt’s Model

Like Garrett’s model, Levelt et al.’s (1999) conceptual network is also a uni-directional staged process, which is shown in Figure 2-3.
Figure 2-3. Fragment of the lexical network (from Levelt, 1989)

Figure 2-3 is Levelt’s lexical network underlying lexical access, which has three stratums—conceptual, lemma and form in serial (Levelt, 1989). Conceptual formation contains two steps—macroplanning and microplanning, representing how the speaker selects information to obtain communicative goals and then orders the complex messages into expressions. And then upon the lemma selection, the appropriate lemma will be activated and so do its related diacritic parameters. For example, verbal features such as aspect, tense, and number will be retrieved simultaneously when activating the verb lemma *escort*. Finally at the form stratum, the morphological form will be retrieved, such as *escort* has two morphemes <escort> and <ing>.

Based on the model, Levelt *et al.* (1999) specified several reasons to explain the phonological influence on semantic errors. First, at the lemma stage, nodes with similar meanings and sounds will also be activated and thus semantic and phonological cohort effects are expected. For example, when the target node is ESCORT, the former effect will coactivate ACCOMPANY which is semantically related and the later effect will activate the phonologically-related node ESCAPE. The semantic activation illustrates why most semantically-related errors involve hyponyms and the phonological effect depicts why most lexical errors have
phonological facilitations. However, semantic and phonological cohort effects only explain errors related in meanings or sounds respectively, rather than mixed errors.

Secondly, Levelt (1983, 1989) distinguished the self-monitor in his production model, to observe if the output is permissible or not. The monitor, also known as post-lexical editor, will only allow legitimate words to be produced. It explains the lexical bias in phonological errors and also describes mixed errors (Levelt et al., 1999). The reason is the editor tends to intercept non-word outputs other than real words. Thus it is easier for mixed errors to go through the monitor than pure semantic errors. It is called ‘monitoring effect’ (Levelt et al., 1999). However, in the principle of simplicity, such mechanism seems unnecessary. Besides, sometimes the editor may be quite inefficient because it still fails to detect a large number of errors (Harley, 1984).

Thirdly, Levelt (1991) submitted mixed errors as ‘environmental errors,’ which means the intrusion of words is caused by factors in the environment, such as distractors, speaker’s attention or a higher feedback from concept to lexical selection. Nevertheless, in the current corpus such environmental intrusions were excluded. Thus this concern will not interfere with the result (see Chapter 3 for more details).

2.2.2 Connectionist Model

Contrary to the uni-directional processing of serial-ordering model, the connectionist model, also known as cascaded or activation spreading model, emphasizes the bi-directional spreading of activation. Thus the top-down and bottom-up activation are both allowed in the processing.

In this section Dell’s spreading activation model (1986) is introduced. As Figure 2-4 exhibits, the lexical network of the model includes three layers—semantic, lexical
and phonological. In this network, the connections of all nodes allow bidirectional spreading between adjacent levels (Dell & O'Seaghdha, 1992).

**Figure 2-4.** Lexical network structure in the spreading activation production model (from Dell, 1992)

The three major operations of spreading activation include spreading, summation, and decay (Dell, 1986). Spreading means the activation level of a node will be greater than zero and thus it can spread to neighboring nodes. Summation means the activation of each node will be amounted. Finally the one with the highest activation will be selected and then decays towards zero. Such process keeps being applied to all selection of nodes at each layer.

As what mentioned before, the connectionist model features the forward and backward interflow of information between layers. It gives a clear illustration of the reason for lexical substitutions, especially ‘mixed errors’—both phonologically- and semantically-related errors (Dell & Reich, 1981; Dell & O’Seaghdha, 1992). To be more clearly, a lexical error is formed when a wrong lexical node is retrieved at the lexical layer. When the target node and the intended node overlap in semantic feature nodes, such as *cat* and *dog*, it will result in semantic-related lexical errors. Likewise, when they have similar phonological nodes, like *rat* and *mat*, it will cause phonological-related errors. Besides, mixed errors, such as *cat* becoming *rat*, will probably be produced due to the bidirectional activation in this model. When a lemma
node is activated, its related semantic and phonological nodes are both retrieved, and the phonological-related nodes will send activation weight back to the lemma level. Finally the node obtaining weight both from the lemma and phonological level is easier to be selected, i.e., mixed errors.

Such view of feedback activation also has the empirical support. For example, Dell and Reich (1981) investigated numbers of identical phonemes of 289 lexical substitutions and found nearly 40% errors involving the identical segment in the first position. Such strong phonological similarity in lexical errors made them argue for the joint effect of phonological and semantic activation. Furthermore, Harley (1984) searched 123 lexical substitutions, taking initial consonant and syllables structures to evaluate phonological similarities. Because both criteria were significant in paradigmatic errors, he asserted these ‘high-level intrusion errors’ were constrained by low-level factors. It contradicted with serial-ordering model in which the message processing should be complete before accessing phonological information. Thus next the comparison between serial-ordering and connectionist model in the literature will be reviewed.

2.2.3 The Comparison between Models

The main difference of the serial-ordering model and the connectionist model is whether each level functions as an independent stage or interactions can flow between levels. This issue has been widely examined in previous researches, primarily in two ways—through psycholinguistics experiments in laboratory or corpus of speech errors in natural settings.

Numerous experimental designs have been conducted to test whether phonological and semantic processes operate independently or have interactive influence. However, the results were unclear or incoherent. First of all, although
Martin et al. (1996) observed phonological similarity in semantic errors of both normals and aphasias through picture-naming tasks, it was unclear whether the interactive activation or post-lexical editor accounted for the data better. Secondly, other studies were not coherent in their conclusions. On the one hand, some supported the serial-ordering models by investigating the time lag between the onset of the distractor word and the subject’s utterance in a picture-naming task (Schriefers et al., 1990; Levelt et al., 1999). They found that effects of phonological distractors were only present in the later stage, and thus favored a discrete model. On the other hand, some researchers, arguing for the connectionist model, suggested that semantically processed but unselected items also undergo phonological activation (Martin, Weisberg & Saffaran, 1989; Peterson & Savoy, 1998; Cutting & Ferreira, 1999; Rouibah, et al., 1999; Morsella & Miozzo, 2002; Rapp & Samuel, 2002). Moreover, Cutting and Ferreira (1999) further conducted the picture-naming task with homophones as stimuli and the result also approved the cascaded model.

Although both groups were sensible, most of their experimental stimuli did not completely cover all the phonological variables when discussing the phonological similarity. For example, the phonologically-related target pairs of Morsella and Miozzo (2002) were only identical in initials and preserved the same numbers of syllables (e.g. BED-bell). Also, Rapp and Samuel (2002) mainly focused on the rhyme effect of their sentence completion tasks.

Therefore, in addition to psychological experiments, the second optional method in scrutinizing two models is through natural corpora. This method intends to investigate more types of phonological similarities once at a time although it surely has some limitations (see 2.1.2 for details).

Studies analyzing natural speech errors aim to examine the way of lexical storage first because it regards lemma retrieval in the processing model. Nevertheless, the
results are inconsistent between two types of models. For the serial-ordering model, many researchers asserted the mental dictionary is organized phonologically, according to their phonemic structures, initials, phonetic features, etc. (Fromkin, 1971; Ryan et al., 1975; Fay & Cutler, 1977; Aitchison & Straf, 1982; Emmorey & Fromkin, 1988). They believed words were retrieved in a cohort fashion (Marslen-Wilson & Welsh, 1978) but it was not sufficient in explaining semantic errors (Emmorey & Fromkin, 1988). Therefore, Levelt (1989) further suggested there were two kinds of internal organization of mental lexicon—based on forms and meanings respectively. It served as evidence for the serial-ordering model in that phonological and semantic retrieval will not exert influence on each other. For the connectionist model, Dell and Reich (1981) suggested that lexicon was represented as a network consisting of linguistic units (features, phonemes, words, etc.) rather than a dictionary-like listing (McClelland & Rumelhart, 1981; Harley, 1984; Stemberger, 1985b; Zhou & Marslen-Wilson, 1995; Roelofs, 1996).

The different ways of lexical storage determine how they interpret semantically- and phonologically-related lexical errors (e.g. mixed errors) in their model. The serial-ordering model argued that mixed errors are due to misselection of neighbors in a phonologically-arranged dictionary, based on a distinctive feature system (Fay & Cutler, 1977). On the contrary, the connectionist model emphasized the spreading activation in the lexical network will retrieve the semantically-related target node. Moreover, the phonological similarities in lexical errors represent there is information leaking between stages of production (Dell & Reich, 1981).

In summary, in order to distinguish these two models, phonological similarities in paradigmatic lexical substitutions become the crucial evidence. An over-all investigation will be made to estimate the intensity of phonological similarity. Its significance will support the idea of information-leaking of connectionist model.
2.3 Linguistic Overview

According to 2.2, phonological similarities in paradigmatic errors are crucial in distinguishing production models. In order to describe it more precisely, this section reviewed some linguistic properties for evaluation, including lexical category, segments, features, syllable structures and tones. Notice that many issues about phonology have excited a series of debate among researchers. Thus the current study only focuses on the topic related to the research questions.

2.3.1 Lexical Categories

Lexical categories concern how words in a language are classified. It regards how to describe each error precisely in the study. Huang et al. (2007) divide words of Mandarin Chinese into four primary lexical categories, namely verbs (V), nouns (N), prepositions (P) and adjectives (A), based on their syntactic behavior. Their difference could be fundamentally distinguished by the proto-categories [N] and [V]. A word will either belong or not belong to proto-N and proto-V, meaning that one category has two-binary features, i.e. [±N] and [±V]. According to it, four kinds of feature combination are shown in the matrix (2).

(2) Feature-based characterization of lexical categories

<table>
<thead>
<tr>
<th>Feature</th>
<th>Category</th>
<th>Noun</th>
<th>Adjective</th>
<th>Preposition</th>
<th>Verb</th>
</tr>
</thead>
<tbody>
<tr>
<td>[N]</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[V]</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>

In (2), it is obvious that nouns are [±N, −V] while verbs are [−N, +V], because nouns often describe things that exist while verbs refer to situations that take place. Notice that classifiers are considered as nouns here. As for adjective [±N, +V] such as (3a), they are assigned [+N] because their semantic objects will be introduced
by *dui* ‘on’, just as nouns. And they behave like verbs since they do not need *shi* ‘am’ in sentences either.

(3) Adjective:

(a) *wo* *dui* tade qushi fejchang *shangxin*

I on his pass.away extremely sad

‘I am extremely sad about his pass away.’

(b) *wo* fejchang *shangxin* tade qushi

I extremely sad his pass-away

(c) *tade* qushi *wo* *shi* fejchang *shangxin*

his pass-away I am extremely sad

Example (3) accounts for the behavior of the adjective *shangxin* ‘sad.’ The evidence of [+N] comes from (3a) and (3b), in which *dui* ‘on’ is necessary to indicate its semantic object *tade qushi* ‘his pass-away.’ And its another feature [−V] is represented in (3c) that *shi* ‘am’ is not required in the sentence.

Prepositions include *guanyu* ‘about,’ *cong* ‘from,’ *gei* ‘to/for,’ *zai* ‘at,’ *xiang* ‘toward,’ *ba* and *bei*. They do not take any nominal subjects and do not function like verbs, thus are assigned [−N, −V].

(4) Preposition:

(a) *guanyu* zhe-jian shi, tamen yijing taolun-guo le

about this-CL issue they already discuss-GUO SFP

‘They have already talked about this issue.’

(b) *tamen* guanyu zhe-jian shi

they about this-CL issue

(c) *women* de huiyi guanyu na-ci shigu

our DE meeting about that-CL accident

Example (4) describes the behavior of the preposition *guanyu* ‘about.’ (4a) is a default position for preposition, i.e., in a pre-subject position. In (4b) it cannot take the
nominal subject *tamen* ‘they’ while in (4c) it cannot be used as a verb directly, proving its feature as \([-N, -V]\) respectively.

### 2.3.2 Segments

Here Mandarin consonants and vowels will be introduced. Consonants are listed in Table 2-1, including 19 consonant phones, 3 palatals and 3 glides.

Table 2-1

<table>
<thead>
<tr>
<th>Place</th>
<th>Bilabial</th>
<th>Labial-dental</th>
<th>Dental</th>
<th>Retroflex</th>
<th>Palatal</th>
<th>Velar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plosive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unaspirated</td>
<td>p</td>
<td>t</td>
<td>k</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plosive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspirated</td>
<td>pʰ</td>
<td>tʰ</td>
<td>kʰ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>f</td>
<td>s</td>
<td>ʂ/ʐ</td>
<td>(ɕ)</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Affricate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unaspirated</td>
<td></td>
<td>ts</td>
<td>tʂ</td>
<td>(tɕ)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affricate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspirated</td>
<td></td>
<td>tʂʰ</td>
<td>tʂʰ</td>
<td>(tɕʰ)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td>ɲ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liquid</td>
<td></td>
<td></td>
<td>l</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glide</td>
<td>(w)</td>
<td></td>
<td>(j) (ʡ)</td>
<td>(w)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2-1 displayed 19 consonant phones in Mandarin Chinese. Both plosives and affricates have a two-way distinction of aspiration ([p/pʰ, t/tʰ, k/kʰ, ts/tsʰ, tʂ/tʂʰ, tɕ/tɕʰ, tɕʰ/tɕʰ]) while the retroflex fricative has a two-way distinction of voicing ([ʂ/ʐ]). Besides, notice that affricates such as /ts/ and /tʂ/ are considered as a single unit rather than a combination of two phonemes.

Glides [j], [ʡ] and [w] are generally considered as allophones of vowels [i], [y] and [u] (Wan, 1999; Duanmu, 2000). However, in our study, glides are considered as single consonants in the word-initial position. For example, the target word [wɔn35]
‘smell’ and the replacing word [wan35] ‘play’ are regarded containing the identical initial consonant [w]. Those glide appearing in the pre-nuclear and post-nuclear position such as [w] in [kwan51] ‘can’ and [kw55] will be defined later.

In Mandarin phonology, the status of palatals [tɕ], [tɕʰ] and [ɕ] is also controversial. Researchers cannot make an agreement that palatals should be independent segments (Cheng, 1973) or allophones of dentals or velars (Wan, 1999; Duanmu, 2000). However, the underlying status of palatals will not influence the analysis because palatals are just considered here as distinct consonants in the surface representation. For example, when [ʂon51] ‘give’ is replaced by [ɕɛn51] ‘present,’ the dental fricative [s] and the palatal fricative [ɕ] are counted as different consonants.

As for Mandarin vowels, their phonemes and their allophones are listed in (5) (Duanmu, 2000; Wan & Jaeger, 2003).

(5) The vowel system in Mandarin

| High | /i/ → [i, i, j] |
| /y/ → [y, ɻ] |
| /u/ → [u, w] |
| Mid | /a/ → [e, e, ə, ə, o, ɔ] |
| Low | /a/ → [a, ɑ] |

Here this 5-vowel system is adopted (Wan & Jaeger, 2003). Glides [ɻ, ɻ, w] are derived from high vowels /i, y, u/ respectively. In addition to that, the high front vowel /i/ has two allophones [i, i] with complementary distributions. Vowels [e, e, ə, ə, o, ɔ] are all variants of the mid vowel /a/ and the low vowel /a/ has two allophones [a, ɑ]. Notice that in the analysis the error such as [ɕja51] ‘down’ replacing [tɕw51] ‘get,’ two vowels [a] and [ɑ] are considered as different vowels although they belong to the identical phoneme /a/. Their difference could be described more clearly in terms of phonetic features, which will be introduced in the next section.
2.3.3 Phonetic Features

In the phonological representation, both consonants and vowels are composed of a bundle of phonetic features which belong to different natural classes. By comparing the number of shared features between segments, their phonological similarities could be evaluated. Take consonants for example, in the phrase *goofing off* uttered as *gooping off*, /f/ and /p/ are quite similar since they share many common phonetic features. However, how to describe their similarity specifically depends on which feature system is used. In this study Wan’s (1999) feature system of Mandarin consonants is adopted, which is developed upon Broecke and Goldstein’s (1980) system of English consonants. The system is produced from the observation of speech errors and thus should be appropriate to describe speech errors. It is presented below.

Table 2-2

<table>
<thead>
<tr>
<th>Phonetic Features of Mandarin Consonants (from Wan, 1999)</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
</tr>
<tr>
<td>P</td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>V</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td>F</td>
</tr>
</tbody>
</table>

Wan (1999) set five parameters to distinguish all features of Mandarin consonants. Their illustrations are given below.

1. *Place of articulation*: this parameter is composed of seven positions for articulation: labial (L), dental (D), velar (V), retroflex (R), palatal (P), labial-palatal (L-P) and labial-velar (L-V).

2. *Nasality*: only three nasals are included—labial /m/, dental /n/ and velar /ŋ/.

3. *Voice onset time*: this dimension has three levels—aspirated (A), voiceless unaspirated (O) and voiced (V). Moreover, all stops and affricates has a two-way
distinction for aspiration (A/O) while fricatives and sonorants contrast in voicing (O/V).

4. Continuancy: all sonorants and fricatives are [+continuant] while other stops and affricates are [−continuant].

5. Frication: all fricatives and affricates are [+frication] while others are [-frication].

Moreover, like consonants, vowels can also be distinguished from the feature description based on five natural classes, namely [±high], [±low], [±back], [±round] and [±RTR] (Wan & Jaeger, 2003). The feature-counting system of vowels is displayed in Table 2-3.

Table 2-3

<p>| | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Low</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Back</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Round</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>RTR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Table 2-3 showed how Mandarin vowels are marked with different feature values. The natural class [±high] and [±low] represent the height of vowels while [±back] emphasizes its position. Besides, [±round] and [±RTR] indicates the manner for articulation. [±Round] refers to the shape of lips while [±RTR] means ‘retracted tongue root,’ which is used to distinguish [ɔ-ɣ], [o-ɔ] and [a-ɑ].

2.3.4 Syllable Structure

This section introduces the traditional syllable structure of Mandarin and how it is utilized in our analysis. The traditional analysis of Mandarin syllable structure is exhibited as (6), including three primary portions—tone, initial and final (Cheng, 1973).
In (6), tones refer to the pitch of syllables. The initial can be a single consonant or a nasal. Finals are the rest segments of the syllable, including medial, nucleus and ending. Medial refers to the glide preceding the vowel. Nucleus are vowels, and endings include vocalic and consonantal, such as glides [j, w] and nasals [n, ñ].

According to this view, the possible syllable structures in Mandarin are listed in (7), in which V represents vowels, C is consonant and G means glides.

(7) Syllable structures:

(a)  V  CV  GV  VC  VG
(b)  u55  lu51  wɔ21  in55  aj51
(c)  ‘dark’ ‘road’ ‘I’ ‘music’ ‘love’

(a)  CVC  CVG  GVC  GVG  CGV  CGVC  CGVG
(b)  man21  taj51  joŋ51  jow51  twɔ55  ljen21  xwaj51
(c)  ‘full’ ‘bring’ ‘use’ ‘again’ ‘many’ ‘face’ ‘bad’

Example (7a) illustrates all combinations of segments in Mandarin. Their variable length ranges from a single V to a maximum of four, CGVC or CGVG. (7b) shows the IPA transcription while (7c) gives the English gloss.

Furthermore, to determine the hierarchical structure of syllables involves a controversial issue, i.e. the status of prenucleus glides. This study adopts the traditional view which assumes glides are grouped with the rime. Such view has the empirical support that the \textit{fanchie (反切)} experiment proved that subjects tended to
classify glides with the rime (Wang & Chang, 2001). Based on this view, the hierarchical structure is provided in (8).

$$\begin{align*}
\text{Example (8) is the internal syllable structure, in which O means onset, R means rime, N is nucleus and C_o stands for codas. Thus in the analysis, onset is equal to the initial segment and the maximum of rhyme includes a glide, vowel and coda.}
\end{align*}$$

### 2.3.5 Tones

Mandarin Chinese has four distinctive tones and one neutral tone. The total pitch range of tones are divided into five levels by Chao (1930), marked with numbers 1 to 5. ‘1’ represents the speaker’s lowest pitch and ‘5’ is the highest. The correspondence of the pitch and tone are as (9) shows.

$$\begin{align*}
\text{(a) high} & \quad \text{rising} & \quad \text{low-falling} & \quad \text{falling} \\
\text{(b) ma55} & \quad \text{ma35} & \quad \text{ma21} & \quad \text{ma51} \\
\text{(c) ‘mother’} & \quad \text{‘hemp’} & \quad \text{‘horse’} & \quad \text{‘scold’}
\end{align*}$$

Example (9) is the representation of four distinctive tones. (9a) illustrates the quality of each tone: the high, rising, low-falling, and falling tone respectively. (9b) marks the number of each tone in the example of syllable [ma]—55, 35, 21, 51 respectively. Their responding gloss is provided in (9c). This example reveals that tones in Mandarin are used to distinguish lexical meanings and thus the syllable [ma] exhibits four different meanings when connecting to four different tones. In other words,

---

1 In Beijing Mandarin T3 should be pronounced as 214, a falling-rising contour. However, it is only produced as a low tone 21 in Taiwan Mandarin (Wan, 1999).
Mandarin tones are specified in the lexicon, unlike the phrasal stress in English which is derived by rules in the surface structure (Cheng, 1973).

2.4 Phonological Effects in Previous Studies

In the previous section, linguistic properties such as lexical category, segment, feature, syllable structure and tone in Mandarin have been reviewed. They function as the measurement system to quantify the phonological similarities in paradigmatic errors. For example, lexical category of errors is represented as category effect. And errors started with the same consonant exhibit the initialness effect. Other effects such as similarity, syllable structure and tone effect will be introduced in series. This section is going to investigate how these effects are described in previous studies and all of them will be examined in Chapter 4.

2.4.1 Category Effect

Category effect means the interacting units in substitutions often preserve the same syntactic category. It was evident in both syntagmatic and paradigmatic lexical substitutions (Nooteboom, 1969; Dell & Reich, 1980; Stemberger, 1985b). Stemberger (1985b) attributed it to the existence of syntactic condition in the slots. In detail, slots generally associated with syntactic features such as [+noun] and [+verb]. The feature will narrow down the access of lexical items and thus increase the possibility to activate items of the same subgroup. Therefore, a noun slot is more available for nouns rather than verbs since targets as nouns will receive greater activation at the semantic level. As a result, the category effect is manifested in lexical errors.
2.4.2 Initialness Effect

The status of initial consonant is crucial because in syntagmatic errors such as *tonal phonology* becoming *fonal phonology*, initials tend to slip more than other parts (Boomer & Laver, 1968; Mackay, 1970; Aitchison & Straf, 1982; Dell, 1986; Shattuck-Hufnagel, 1986; Stemberger & Treiman, 1986; Stemberger, 1989). Besides, Dell and Juliano (1996) reported about 80% consonant errors misordered with initial onsets. Wan (2007b) also found in 1375 errors, 67.67% had target and source in syllable initial position.

The effect in syntagmatic errors received two possible explanations: One was proposed by Dell (1986) that initial consonants may activate stronger connection weight from higher levels and thus tended to slip. The other was from Mackay (1972) and Shattuck-Hufnagel (1986) who attributed it to the hierarchical syllable structure that a CVC syllable should contain two syllabic constituents—an onset and a rhyme. The vulnerability of initials implied the distinctive status of onsets. In such a view, initial slips could be a consequence of sequential retrieval (Dell & Juliano, 1996).

In addition to syntagmatic errors, the initialness effect was also found in paradigmatic errors. For example, the interacting units tended to have the identical initial consonant (Dell & Reich, 1981: 40%; Harley, 1984: 24.4%; Wan, 1999: 20.1%). The significance of initialness effect strengthens the phonological similarity in paradigmatic lexical errors. Therefore, the joint effect of phonological and semantic activation in word selection implies information may leak between levels. It means the decision-making in one stage will probably be influenced by other stages. (Dell & Reich, 1981).
2.4.3 Similarity Effect

Similarity effect means that phonetically-similar segments tend to interchange with one another. It was significant in previous studies (Nooteboom, 1969; Mackay, 1970; Fromkin, 1971; Fay & Cutler, 1977; Shattuck-Hufnagel, 1979; Levitt & Healy, 1985; Stemberger, 1985b, 1989; Wan, 2007b).

The effect was established in both interacting units and their environment. First, in the exchange error reading list becoming leading rist, two initial consonants [l] and [r] share many but one phonological features: [±lateral]. Fay and Cutler (1977) observed 156 malapropisms and found nearly 25 percent differed in only one feature. Secondly, in left hemisphere becoming left hemisphere, the exchanging consonants have similar environment, where these two consonants are both followed by the vowel [ɛ] (Nooteboom, 1969; Mackay, 1970). Both dimensions should be considered in the measurement of similarity effect.

The similarity effect reflects several issues about word production. First, in terms of lexical storage, 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 target next to the intended unit tended to be selected and thus generated lexical errors. However, their argument is limited in describing the similarity effect rather than interpreting all other phonological effects. Secondly, in the process of word activation, the significance of similarity effect illustrates the cognitive status of features. For example, Wan (1999) examined speech errors of Mandarin Chinese and found that feature values can move or be replaced as independent units but not as robust as segments. Thus she concluded features also have a cognitive status in speech production, only that segments are primary planning units (Fromkin, 1971). Therefore, the value of similarity effect should be compared with that of initialness effect, to see which one is primary in speech production.
2.4.4 Rhyme Effect

Rhyme effect intends to examine if interacting units of lexical errors share their rhymes, including CV, VC and CVC.

The effect 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. In order to test the hypothesis, two things need to be approved—One is onsets should be more likely to err than codas, which has been summarized in 2.4.2. The other is VC sequences should occur more frequently as error unit than CV sequences because VC corresponds to the sub-syllabic constituent of syllable structure. In previous researches, VC sequences were frequently preserved in phonological errors (Nooteboom, 1969; Mackay, 1970; Shattuck-Hufnagel, 1979; Rapp & Samuel, 2002; Wan, 2007b—VC: 4.78%, CV: 0.08%). Therefore, this study aims to examine if such rhyme effect also exists in paradigmatic lexical errors and to determine it is CV, VC or CVC as the primary erroneous unit. If the rhyme effect (VC) is significant, the hierarchy of syllable structure is further approved.

2.4.5 Syllable Structure Effect

The construction of syllable structure is strongly respected in phonological errors, especially single-segment substitutions. It is called Syllabic Similarity Phenomenon that the interacting segments usually occupy the identical position in the syllable, i.e. onsets replace onsets, nucleus replace nucleus and codas substitute codas (Boomer & Laver, 1968; Mackay, 1970: 98%; Fromkin, 1971; Shattuck-Hufnagel, 1979; Wan, 2007b: 99.22%). Boomer & Laver (1968) proposed one explanation that the syllabic position is constrained by the manner of articulating the phoneme.
In addition, syllable structures are often preserved in paradigmatic lexical errors. For example, Harley (1984) claimed in 30 paradigmatic errors, 21 of them contained the same syllable pattern. Such similarity of syllable structures between interacting words is illustrated in Dell’s (1988) model. He proposed that each word node in the lexical network connects to a word-shape ‘header’ node with the pattern CVC or CV and then it links to other phoneme categories such as Ci (initial), V and Cf (final). Finally each phoneme category node connects back to all possible phoneme nodes like /b, d…/. In this framework, the number of activated phonemes is determined by the word-shape header, called ‘categorically triggered selection.’ Nodes with the same syllable pattern are activated and finally the highest one will be selected.

However, in studies of Mandarin, the effect was not so significant. Wan (1999) examined 224 paradigmatic errors and reported that only 12% contained the same syllable structure, which was lower than the chance level (12.5%). She thus concluded syllable structures are assigned later in phonological encoding rather than being part of phonological organization of the lexicon.

Therefore, the effect will be reevaluated in this study to determine the role of syllable structural nodes in the processing model.

2.4.6 Tone Effect

The position of tone in production models did not get reconciled in previous researches. Some pronounced it as part of underlying lexical form while some argued it was not represented until the phonetic stage.

On the one hand, Beckman (1986) asserted tones linked more closely with lexical items and Wan (2007a) further proved it by postulating the fact that lexical substitutions involving the same tone were greatly more than that of segmental
substitutions (lexical: 56%; segmental: 29%). She thus concluded tones were part of phonological organization of the lexicon.

On the other hand, Chen (1999) argued that tones were suprasegmentals in their underlying forms and only became associated with rhymes or vowels in the phonetic configurations. The conclusion was drawn from both natural corpus and experimental tasks. In examining the error corpus, Chen (1999) argued that true tone errors were relatively rare in Mandarin and thus tones were not processed similarly as segments. Within the implicit priming paradigm, he found that the syllable-alone prime produced some priming effects while the tone-alone prime produced none (Chen & Dell, 2002). It proved that the syllable without tone was an independent planning unit while the tone alone did not behave as a unit like the syllable does. Therefore in his opinions, tones were part of the phonological frame rather than lexical organization (Chen, 1999).

Between these disparate views, tone effect in paradigmatic errors may help clarify the status of tones in word-production process. If it is significant, tones may be part of lexical organization in Mandarin Chinese. Besides, if the intensity of tone and rhyme effect are not comparable, it may reject Chen’s view (1999) that tones are associated with rhymes in phonetic configurations.

2.5 Summary and Research Hypotheses

In 2.1 the definition and application of slips of the tongue are introduced. Through speech errors, the serial-ordering and connectionist model are constructed. Their theoretical arguments are illustrated in 2.2 and their respective empirical evidences are shown in 2.4.

Researchers arguing for the connectionist model indicated that lexical substitution with phonological similarities occurred more frequently than chance because the
phonological activation of unselected lexical nodes will leak from the phonological stage back to the lemma stage (Dell & Reich, 1981; Harley, 1984). However, those for the serial-ordering model advocated that mixed errors were due to the double malfunction of separate levels, which may not be favored if the proportion of mixed errors amounts higher than chance. If mixed errors take the larger proportion in the corpus, it might support the connectionist model.

Based on it, the phonological similarities in paradigmatic errors turn to the important evidence for the connectionist model. First, the category effect is investigated since it substantiates the nature of paradigmatic errors, which are different from those phonological errors with lexical bias. Besides, it will ensure that each syllable-position slot only accepts its corresponding content.

Secondly, the significance of initialness effect will imply lexicons are processed by initial sounds. Initials are considered as ‘retrieval cues’ in phrasal processing (Garrett, 1988). In previous studies, they tend to slip in phonological errors (e.g., MacKay, 1970; Aitchison & Straf, 1982; Dell, 1986; Stemberger, 1989; Wan, 2007b) and to remain identical in semantic substitutions (Dell & Reich, 1981; Harley, 1984; Wan, 1999). In sum, it could be the consequence of sequential retrieval in the processing model (Dell & Sevald, 1994; Dell & Juliano, 1996).

Thirdly, the intensity of similarity effect has been proven in phonological errors (e.g., Fay & Cutler, 1977; Shattuck-Hufnagel, 1979; Levitt & Healy, 1985; Wan, 2007b). Thus in the analysis of paradigmatic lexical errors, the presence of similarity effect will suggest that phonetic features are already available during the lemma-selection stage. Conversely, its absence indicates phonetic features are in the bottom of phonological encoding, not involving in the feedback activation.

Fourthly, the rhyme effect plus initialness effect in Mandarin will approve Syllable Structure Hypothesis (Hockett, 1967) that a hierarchical structure of CVC
sequence is composed of the onset (C) and the rhyme (VC) (e.g., Boomer & Laver, 1968; Mackay, 1970; Fromkin, 1971; Shattuck-Hufnagel, 1979; Wan, 2007b). However, if they are not significant, it means Mandarin does not support such division of syllables.

Fifthly, if the structures of interacting syllables are similar, it may prove that the processing of Mandarin also has various word-shape nodes such as CVC or CV in the lexical network and all participate in the feedback toward lemma selection (Dell, 1988). On the contrary, the inconsistent syllable structures may imply modifications are needed in Dell’s (1988) model.

Finally, the status of tones is controversial in previous studies. Wan (2007a) reported it was represented as the underlying lexical organization while Chen (1999) argued it was part of metrical frame. Therefore, if tone effect is significant, tones may be part of lexical organization in Mandarin. If it is not, it turns out that tones are more superficial in phonological organization.

Following the reasoning, seven research hypotheses are provided in details below.

1. It is supposed that in paradigmatic lexical errors, lexical substitutions which are both phonologically- and semantically-related (i.e., mixed error) will be predominantly more than those pure phonologically- or semantically-related.

2. The intended utterance and the target word should respect their lexical category, i.e. nouns only replace nouns and verbs only replace verbs, etc.

3. In terms of the phonological similarities in paradigmatic errors, they should display the initialness effect. In other words, most interacting units are supposed to contain the identical initial consonant.

4. The intended utterance and the target word are believed to show the similarity effect. It means both the interacting consonants and vowels should share more phonological features than average.
(5) The paradigmatic errors should have the rhyme effect, and as for the error unit, VC or CVC is supposed to participate more than CV.

(6) The corpus should display the syllable structure effect in a significant sense. That is, most intended utterances and the targets will contain the same syllable structure.

(7) In the paradigmatic lexical errors, most of them should contain the same tone.

As what has been mentioned, the first hypothesis may turn to be the evidence for the connectionist model. As for others, they may facilitate the phonological similarities among paradigmatic errors and also have some reflections on the word production process and lexical storage in Mandarin. All of them will be examined in Chapter 4 in series.