

SEMANTIC RELATIONSHIPS IN MANDARIN SPEECH ERRORS*

I-Ping Wan¹ and Jen Ting²

¹*National Chengchi University*

²*National Taiwan Normal University*

ABSTRACT

This paper investigates various patterns in a corpus of naturally-occurring lexical substitution speech errors to show how Mandarin speakers produce such lexical items and how they relate to the lexical-semantic relationships in Mandarin lexicon. Target-error pairs often share similar semantic features or are semantically-related associates and can be categorized into general taxonomies of semantic relatedness. A coordinate relationship is the most common type, followed by association, and contrastive is the least common type. These findings show that two lexical items are related by sharing a number of semantic features or sharing values on these features to the same level of specificity in the context. Lexical errors in Mandarin confirm some general findings in other relevant cross-linguistic studies, suggesting that the semantic links between two lemmas in the lexicon have caused the error. The error distributions in the Mandarin corpus suggest that lexical selection occurs independently in models of lexical production. The data confirm the description of the formulation stage in the context of psycholinguistic models of sentence production (Garrett, 1975, 1982, 1984; Levelt, 1989; Dell, 1986, 1988).

Key words: lexical substitution errors, semantic relations, speech errors, Mandarin

* We are very grateful for the following scholars for discussions on some of the issues: Prof. M. T. Chang, Prof. K. Chui, Prof. S. Gong, Prof. R. R. Huang, Prof. J. Jaeger, Prof. H. Lai, Prof. Y. Li, Prof. Y. Lin, Prof. C. Liu, Dr. J. Tsai, and Prof. H. I. Wu. We also like to thank all the research assistants for organizing the data structure: P. Chen, Y. Cheng, T. Fu, H. Hsu, K. Lee, S. Lin, K. Lu, K. J. Lu, and H. Wang. Errors and infelicities which remain are our own. Research reported in this paper has been funded to the first author by NSC grant, NSC 98-2410-H-004-103-MY2, in Taiwan.

1. INTRODUCTION

Many useful sources such as speech errors, aphasic studies, or psycholinguistic experiments involving priming or word association tasks serve as important evidence for the semantic organization of the lexicon (Clark, 1970; Levelt, 1989; Kess, 1992; Martin et al., 1994; Alario and Cohen, 2004). All these studies examining the effects of semantic relatedness of words account for patterns of behavior. For instance, evidence supporting the separation of concept and lemmas in representation and processing comes from lexical errors in a natural speech-error corpus (Levelt, 1989) and a psycholinguistic experiment (Levelt, Roelofs and Meyer, 1999).

Speech errors were collected and analyzed by many researchers between the 1970s and 1980s. Using error data to test certain linguistic hypotheses is a long tradition, and these errors have been taken as important evidence regarding the units and structures involved in the cognitive representation of language. Many theories of speech production planning or phonological theories have been based mainly on speech error data (e.g., Fromkin, 1973, 1983, 1988; Garrett, 1975, 1980, 1984, 1993; Dell and Reich, 1980, 1981; Stemberger, 1985, 1993; Dell, 1986; Levelt, 1989; Dell, Juliano and Govindjee, 1993; Bock and Levelt, 1994).

The majority of speech-error research in Germanic languages has been done in relation to sound structures, phonological systems and even speech planning models (Fromkin, 1973, 1980; Shattuck-Hufnagel, 1979; Cutler, 1982; Stemberger, 1983; Berg, 1987). Using such error patterns can shed light on the architecture and mechanisms of a language production system. A few studies have used natural speech-error corpora to examine the lexical-semantic relationships in the lexicon, and set up general taxonomies of semantic relatedness, which typically involve 3-5 categories (Hotopf, 1980 for English; Levelt, 1989 for English; Wijnen, 1992 for English; Jaeger and Wilkins, 1993 for English; Arnaud, 1999 for French; Harley and MacAndrew, 2001 for English; Jaeger, 2005 for English adults and children). An exception is Jaeger and Wilkins, who proposed 7 categories to capture the similar patterns in semantic relationships shared by the target-error pairs in the dataset. In all the studies, a similar pattern can be observed in the extent to which target-

error pairs often either share similar semantic features or are semantically-related associates such as co-hyponyms (i.e., words from the same semantic field), antonyms (i.e., words of semantic opposition) or near-synonyms.¹

In the context of psycholinguistic models of speech production planning process, it is generally accepted that lemmas in the lexicon are the abstract representations of the morpho-syntactic and semantic information relevant to the meaning and function of lexical items; lemmas are associated with relevant concepts, and while the most appropriate lexical items are being activated by the message representation, lexical items which are semantically related to them will also be activated in the appropriate form representation (e.g., Garrett, 1975, 1982, 1984; Levelt, 1989; Bock and Levelt, 1994). Even in the interactive activation model of production developed by Dell (1986), the lexical nodes are closely linked with semantic features, and during the retrieval process, the operations of the lexical network can spread the semantically related word to the other word. Therefore, the two types of models generally hold a similar view regarding lexical retrieval in language production although the former deals with a unidirectional, top-down process, and the latter deals with the interactive, bi-directional, both top-down and bottom-up processes. The difference between these two models is that the former one deals with serially ordered elements by means of fixed stages specifying what levels of representation and what kind of computations can occur at each stage, but the latter one explains better the interaction between lexical and phonological elements (i.e., malapropisms).

Speech errors can occur at any stage of the speech production planning process (Garrett 1975, 1982, 1984; Dell, 1984, 1986; Levelt, 1989). Errors were classified as lexical if the unit involved was a meaningful morpheme or word. Lexical errors occur when the 'lemmas' of candidate lexical items that are semantically appropriate for the message are activated, and lexical errors can also be caused by conceptual intrusion. Levelt (1989) found that lexical blend errors occur when the same or very similar concepts are activated, and the concepts activate competing lemmas from the lemma lexicon, suggesting that target-error word pairs in lexical blend errors have

¹ Throughout the whole paper, the term "features" are used in a general way to indicate any semantic properties that are represented in lexical entries.

conceptual-semantic relationships with each other, whereas lexical substitution errors occur when the concept to be expressed activates an appropriate lemma, but then this lemma feeds to another lemma in the lexicon which has lexical-semantic links to the intended lemma in that the two lemmas share semantic featural knowledge in the lexicon. Similarly, in Dell's model of lexical retrieval in language production (1988), lexical errors occur due to the priming of semantically related competitors along with the target node, and since lexical nodes are closely linked with semantic features, the erroneously lexical selection occurs when the word nodes share similar semantic relatedness. These two types of models both show how lexical substitution errors in natural speech-error corpus can shed great light on word pairs in terms of their semantic relationships.

In Mandarin, a number of studies have used corpus linguistic data with computation skills to examine the word meanings that have a certain kind of semantic relations in support of a language-independent theory of lexical semantics, and some researchers have also accessed the database to verify the separated levels of representation between concepts and lemma (e.g., Ahrens, 1994; Tsai et al., 1998; Huang et al., 1998; Seto, 1999; Huang et al., 2000; Myers and Gong, 2002; Huang et al., 2003; Liu et al., 2005; Gong et al., 2008; Liu and Chiang, 2008; Liu et al., 2008; Ahrens, 2012; Chang et al., 2015; Yong, 2015). It is thus important to discover the regularities in the target-error pairs and to interpret the error distribution in terms of the lexical representation and the organization of the lexicon. In this study, we will look at a detailed taxonomy of lexical-semantic relationship between the target and error words. We aim to find out when two lexical items are substituted for each other, whether or not the word pairs are semantically-related, and to what degree they are related. Thus, this paper will be concerned with any kind of semantically-related associates involved in the word pairs.

The paper is organized as follows: In section two, we will discuss the relevant findings in English and other languages. In the third section, we will present in detail the methodology for the collection, classification, and analysis of speech errors in this study. Section four presents our findings, in which we catalog all the various influences and the rank-ordering caused by a semantic relationship between the target and error words, and explains the data in terms of the content and organization of

the lexicon. Comparison of whole word substitution errors in Mandarin and in other languages will also be made. Finally, section five concludes this paper.

2. LITERATURE REVIEW

In the past, a psycholinguistic experiment such as word association tasks was generally used to investigate the production of words with the aim of understanding associative links of varying strengths in the mental network. Clark (1970) in the word association task found that words in antonymic relationships are considered the closest semantic relationships since they are the most commonly produced association responses in word association tasks, followed by words in coordinate relationships, then words in superordinate/subordinate relationships, and then words in near-synonyms and words in association relationships. Kess (1992) stated that word association responses classified in respect to semantic types involve paradigmatic/syntagmatic responses or subordinate/superordinate responses; paradigmatic responses involve members of the same part of speech often including near-synonyms (i.e., thin/skinny) or antonyms (i.e., black/white), and syntagmatic responses involve different parts of speech (i.e., dig/hole). Subordinate and superordinate responses involve the members of the same taxonomy, and superordinate responses are more common than subordinate ones.

In cross-linguistic studies, there are very few reports of lexical substitution errors (Coltheart, 1980; Hotopf, 1980; Levelt, 1989; Wijnen, 1992; Arnaud, 1999; Harley and MacAndrew, 2001). Two to five taxonomies of lexical-semantic categories between target and error word pairs are usually drawn in these studies. The semantic relationships between the target and error words are generally classified as five categories: antonyms, hyponyms, near-synonyms, coordinates and associates (Hotopf, 1980; Levelt, 1989; Wijnen, 1992; Arnaud, 1999; Harley and MacAndrew, 2001). It is stated that the coordinate errors usually occur when speakers are required to supply a relatively more abstract concept; hyponyms or subsumative errors occur when speakers have a greater understanding of taxonomic relationships and the speakers

are aware of lexical entities which can be referred to with a variety of levels of specificity. Synonym errors may suggest that the speakers have more words with overlapping meanings eligible for misselection. Associated errors mean that the speakers have superior vocabulary items and thus a greater range of associations is possible. Antonyms or opposed errors mean that the speakers have words with contrastive meanings that are also possible for misselection.

Coltheart (1980) divided errors into two types to the extent that the target-error words share the same semantic features and they involve some kind of association, but he did not present any numbers regarding the proportion of the categories in the data. Hotopf (1980) reported on a corpus of 224 semantic speech errors in English, the data of which included both paradigmatic and syntagmatic content word substitutions. According to him, regarding semantic substitution errors, co-hyponyms (same superordinate words), occur more often than complementary-antonymous-or-converse (errors), followed by hyponymous relationships (more distant superordinate words). However, he did not find lexical substitution errors involving synonymous relationship. Levelt (1989), despite not providing a figure presenting various categories, agreed that lexical substitution errors most often involve the relationships of antonyms and co-hyponyms, followed by co-hyponyms with a more remote relationship. He also pointed out that there are no superordinate and synonymic errors in substitutions because it is difficult for a listener to recognize the error. Wijnen (1992) looked at 47 paradigmatic errors in the "London-Lund" corpus in English (Garnham et al., 1982) and classified the data into coordination (including co-hyponyms), synonyms, antonyms, associatives (including distant co-hyponyms or metonyms), and others (including collocational items). He found that the majority of the errors are in coordination, while both synonyms and antonyms are very rare. Arnaud (1999) collected 332 lexical substitution errors in French, and he found that most word pairs involved in substitution errors are coordinates (including co-hyponyms or co-paronyms) and antonyms. Synonyms or superordinates/subordinates do not occur in his dataset. Harley and MacAndrew (2001) collected 784 lexical substitution errors in English, which are paradigmatic. They defined shared-feature errors as being from the same semantic field or hierarchically related, involving the

following sub-categories as synonyms, antonyms, superordinates, subordinates and coordinates, and having an association between the target-error words (cf., metonym, Jaeger, 2005). In the data, the majority of errors were synonyms, superordinates, subordinates and coordinates, which were combined in the same group, the next most common were antonyms, and the least common were associations. Jaeger (2005) collected 133 adult errors involving paradigmatic orders, syntagmatic orders, and word blends, and in paradigmatic substitution errors involving content words, she found that the largest percentage of errors involves coordinate types, followed by associated types. The other three types of relationships, subsumative, synonyms, and contrastive, all fall into lower percentages.

Although some of these researchers might have a smaller data set than the others, and the figure might not be reliable to make a comparison, some interesting results have been generalized regardless of the fact that different criteria were drawn. For semantic relationships between target and errors words, the coordinate-type relationships are the most common. The next common type is that of association, except for Hotopf (1980), who observed that antonyms are the next common in his dataset. Basically, the speech-error studies seem to show different results as opposed to Clark's (1970) early findings from a word-association task.

The afore-mentioned studies, except for Jaeger (2005), all presented taxonomies of semantic relatedness only, and they did not provide factors which led to the distribution or rank-ordering based on certain types of semantic features although lexical substitution errors often involve close relationships. Jaeger (2005) proposed numerous factors that might explain the causes of semantic relationships. First, when a speaker is planning an utterance and planning to activate the set of lexical items relevant to the concepts in the proposition, this activation is most likely to spread to other lexical items which are related to the target lexical items by sharing a number of semantic features and sharing values on these features to the same level of specificity in the context. Therefore, coordinate and associative errors are more common. Second, contrastive or opposed errors are less common, since in their corpus, most of the contrastive errors involved either adjectives or adverbs, which do not occur as frequently as nouns, and since this is the type of lexical item that is most likely to have

an antonym or converse in the lexicon, several contrastive errors involved some sort of environmental or discourse influence. As Levelt (1989) pointed out, true synonymic lexical substitution errors are unlikely to be detected, since the hearer would have no way of knowing that the speaker did not intend the spoken word, unless a correction was made. In addition, Jaeger (2005) found that subsumative errors are rare as well, since both the higher-level term (superordinate, set, whole) and lower-level term (subordinate, member, part) would have the same reference, and so would be propositionally equivalent. The hearer would not easily detect it as an error since the substituted word would have the same truth value as the target word. Jaeger's study (2005) is the only one presenting factors that might explain the results. In terms of psycholinguistic models, either Levelt's, Garrett's or Dell's model, pertaining to lexical selection in production, might also confirm such points in that one word that is more semantically related to the other will have a particular word node with a stronger connection. In general, the stability of semantic processes contributes significantly to the probability of accurate lexical selection.

In summary, evidence from cross-linguistic studies seems to show some disagreement as to whether there is a universal rank-order pattern in the semantic features or associations. In our study, we found that the figure in the Mandarin speech-error natural corpus is sufficient to examine these questions thoroughly, and we are hoping to find the lexical-semantic relationships in target-error word pairs and to examine the organization of the lexicon as well as the taxonomy of semantic relationships.

3. METHODOLOGY

The speech errors investigated in this study were obtained from a corpus with over 8000 errors produced by over 100 native speakers of Taiwan Mandarin collected by the first author in a naturalistic setting. The corpus contained largely phonological errors (64%) and the remainder involved lexical errors, morphological errors, phrasal errors, and syntactic errors (Wan, 2007a). The conversations and discourse were tape-recorded (between 1995 and 1999) or digitally recorded (between 2000 and 2009), and detailed Chinese texts were transcribed with the intended target and

the environmental context noted. In an attempt to avoid bias toward particular types of errors, an analysis was made independently by an assistant and the first author when each potential lexical error occurred. Notes were compared, and inconsistencies were resolved by negotiation. Although many researchers have criticized the use of naturalistic speech error corpora due to concerns about reliability and bias, the data involved only semantic errors. Sound errors might be biased or misdetected, but semantic errors are more easily targeted and noticed (see Harley and MacAndrew 2001 for more discussions).

Distinguishing parts of speech or subtypes of semantic relationships regarding shared-feature errors or associative errors in Mandarin is always difficult. To classify the errors, the authors used a scheme similar to that used in Jaeger (2005), but with a few modifications.

Phonological errors can be easily distinguished from true lexical substitution errors in that lexical errors nearly always preserve their original lexical categories, and are usually semantically related to the intended word. Phonological errors, on the other hand, frequently violate their lexical category and have no semantic relationship to the target word; thus they typically produce ungrammatical or meaningless utterances; most importantly, the majority of phonological errors involve source and target units in the utterance (Wan 1999, 2007a, b). In earlier studies, Nootboom (1973) found that 60% of speech errors involve phonological types in Dutch. In English, the figures are 54.4% for phonological errors, and 24.9% for lexical errors in Stemberger's (1989) corpus. In general, phonological errors are more common than lexical errors. Wan (2007a, 2007b) found that phonological errors are 1.8 times, and 7 times more liable to occur than lexical errors, respectively, in her natural error corpus and aphasic speech corpus in Mandarin.

This study is concerned with the semantic relationships between the target and error words in word substitution speech errors. Following Harley and MacAndrew (2001), we only focus on content word substitutions with the exclusion of function words because function words have been shown to behave very differently from content words (Garrett, 1980). Before we move on to explicate the types of the semantic relationships under investigation, it is crucial to define what counts as a content word in this study, for which we adopt Packard's (2000) theory of

word classification in Chinese. Based on the criteria of free-bound and content-function, Packard identifies four types of morpheme. A morpheme that is free and is a function morpheme ([+free, +function]), is a function word, exemplified by the modification particle *de*, the sentence final particle *le*, and the conjunction particle *he* 'and'. A morpheme that is free and is a content morpheme ([+free, -function]) is a root word, illustrated by *bing* 'ice', *shan* 'mountain', and *ma* 'horse'. A morpheme that is bound and is content ([-free, -function]) is a bound root, examples of which include *fang* 'house', *xiang* 'rubber', and *jian* 'perceive'. A morpheme that is bound and is grammatical ([-free, +function]) is an affix, examples of which include word-forming affixes such as the verbalizing affix *hua*, and grammatical affixes such as the perfective aspect marker *le*. Among the four morpheme types, bound roots and grammatical affixes cannot stand alone and thus they are not words. Function words are also excluded, so only root words are content words.

In addition, content words in Chinese may involve "complex words" under Packard's classification. According to him, except for the function words, the other three types of morpheme can freely combine, via morphological processes, to form complex words, yielding compound words, bound root words, derived words, and grammatical words. Compound words consist of two root words, examples being *bing-shan* 'iceberg' and *ma-lu* 'street'. Bound root words consist of a root word plus a bound root or two bound roots, illustrated by *xiang-pi* 'rubber' and *ting-jian* 'hear'. Derived words are bound roots or root words plus word-forming affixes, such as *fang-zi* 'house' and *dian-hua* 'electrify'. Grammatical words are root words plus grammatical affixes, such as *zou-le* 'went' and *xuesheng-men* 'students'. Among the four types of complex words, grammatical words were not included in this study because the formation of such words involves grammatical affixes, which have a [+function] feature.

To classify the errors in terms of the semantic relationships between the target and error words, we used a scheme similar to that used in Jaeger and Wilkins (1993) and Jaeger (2005), and grouped the semantic relationships into five sub-groupings: coordinate, subsumative,

contrastive, synonymous, and associated relationships. In the following, we explicate the classification system, based on Jaeger (2005).²

To begin with, the group of coordinate relationships includes co-hyponyms, co-members of a set and co-paronyms. The two words sharing a coordinate relationship have an equivalent status or the same level of specificity: they may be members of a taxonomy that share the same superordinate, i.e., being co-hyponyms, illustrated by the word pair ‘apple-orange’, where the two words share the same superordinate ‘fruit’; members that share an affiliation in a set, i.e., being co-members of a set, illustrated by the word pair ‘blanket-pillow’, where the two words belong to the same frame which refers to items typically found on a bed and used for resting and or sleeping; or words that label parts of the same whole, i.e., being co-paronyms, illustrated by the word pair ‘finger-thumb’, where the word ‘finger’ and the word ‘thumb’ both refer to part of the hand.³

The second group, that of subsumative relationships, includes superordinate/subordinate, set/member, and whole/part relationships. According to Jaeger (2005), in this group, one of the words denotes elements that are subsumed by elements denoted by the other word. That is, one of the words may denote elements that are subordinate to, are a member of, or are a part of, elements denoted by the other word. The three cases are illustrated by the word pairs ‘do-show’, ‘zoo-donkey’, and ‘thigh-leg’, respectively.

² In this paper, we exclude the errors without semantic relationships and focus more on those with such relationships. Similar methodology has been adopted in the relevant research. Take the study of the thematic relations between a verb and its object as an example. It is not clear what the thematic relation is between an intransitive verb and its cognate object as in (i), but that does not affect the identification of clear thematic relations carried by other instances of objects such as theme, patient, source, etc. (see Gruber 1976, among other).

- (i) a. Sam smiled an evil smile.
b. Jerry danced a merry dance.

³ According to Jaeger (2005), the relationship of co-members of a set is more concrete than that of co-hyponyms. This is because the concept of the ‘set’ involved in this semantic relation is more concrete than the superordinate that is involved in the relation of co-hyponyms.

The third group, that of contrastive relationships, includes binary antonyms, gradable antonyms and converses. Binary antonyms have a relation in which one of the words denotes one property or quality in an entity or event, and the other word denotes the absence of the property or quality, as illustrated by the word pair ‘together-apart’. Gradable antonyms have a relation in which the two words mark the two ends of a perceived continuum, and thus the negation of one word does not entail the confirmation of the other word, as illustrated by the word pair ‘slow-fast’. Converse antonyms have a relation where the two words describe basically the same state or event, and the difference between them only lies in the thematic order of presentation. Take the word pair ‘teach-learn’ as an example: when one teaches something to someone else, it entails that person learns from him/her. It is necessary to note that a target-error word pair may involve “a vaguer notion of contrast which crosscuts other classes, particularly the classes of co-hyponyms and co-members of a set” (Jaeger 2005:325). Following Jaeger’s procedures, we classify such cases like ‘Daddy’ and ‘Mommy’ as coordinates, a semantic relation primarily manifested, rather than canonical contrastives.

The fourth group is the near-synonym set where the target and error words in the word pairs are essentially referentially or propositionally equivalent. An example provided by Jaeger (2005) is the word pair ‘bring-get’, where both words describe activities of transporting objects to someone. According to Jaeger, many instances under her analysis are actually quasi-synonyms, differing, for example, in the level of formality, such as copier/pomper, a neutral vs. slang way to refer to someone who cheats, or in different requirement of selectional restrictions, such as ripe/done. Given this consideration, we rename Jaeger’s class of synonyms as near-synonyms.

The fifth group called associated relationships, includes metonyms, connotation, causal, and shared criterial features. In these cases, the words in the word pair have a relationship which is associated or more loosely connected from a conceptual point of view. A metonymic relation is a relation between two words which refer to things that tend to be either spatially contiguous, as illustrated by the word pair *diaper-light* under a diaper-changing context, or which refer to notions which tend to appear together when discussing a particular scene, illustrated by the word pair

Daisy-Disney, where *Disney* substitutes for *Daisy* in the phrase *Daisy duck* as *Disney* and *Daisy duck* are often discussed together. Connotation labels a relation between two words which are strongly related by connotation, as illustrated by the word pairs *lullaby-love* and *special-fragile*.⁴ Causal labels a relation between two words which are associated by a link of causality, illustrated by the word pair *hurt-hit* as *hit* causes *hurt*. Lastly, the type of shared criteria features describe a relation where two words are semantically related, but do not fall neatly into any of the other groups. Examples provided by Jaeger (2005) include word pairs like *again-yet* and *high-fly*, as *yet* and *again* both indicate temporality and *fly* and *high* both entail the feature of upness.

Adopting the procedures as in Jaeger (2005), we assigned each error pair to the closest semantic relations in the following ranking order: (i) most closely related: coordinates, synonyms, and contrastives; (ii) second most closely related: subsumatives; (iii) third most closely related: the three associated types including metonyms, connotation and causal, and (iv) the most loosely related category: the category of shared criteria features. Thus, although *zuichuen* is 'lip' → *shetou* 'tongue' may involve a metonymic relation (given that the two are spatially contiguous), the word pair is assigned as having a coordinate relation (or more specifically as co-paronyms) under our analysis since on the ranking scale a coordinate relation is a more specific semantic relation than a metonymic relation.

Also, as mentioned earlier, when a word pair shows a seeming opposition relation which crosscuts other classes, it is not classified as an instance of canonical contrastives, but rather as an instance of the primarily manifested semantic relation such as co-hyponyms and co-members of a set. Take, as examples, *zaoshang* 'morning' → *xiawu* 'afternoon' and *nan* 'male' → *nu* 'female'. They are analyzed as co-hyponyms in this study.

⁴ According to Jaeger (2005:327), "two words related strongly by connotation can be substituted for each other." Based on this, Jaeger (2005:328) points out that if something is 'fragile' it is treated as 'special', i.e., it entails that it is 'special'. The connotation is rendered possible under a context where a new delicate plate is talked about. We appreciate the comment from one of the anonymous reviewers for urging us to clarify the role played by context in the occurrence of connotative errors.

Another point worth mentioning in our classification is concerned with the criterion of a whole word. A word pair *li* ‘fence’ → *yan* ‘eave’ may at first sight not involve whole word substitution. However, if we consider the word pair occurs in the idiomatic phrase, i.e. *ji ren li xia* ‘depend on other people for a living’ → *ji ren yan xia* ‘meaningless’, we then find that both *li* ‘fence’ and *yan* ‘eave’ are whole words rather than morphemes of a word, because the context is Classical Chinese where monosyllabic morphemes serve as words. The same reasoning applies to word pairs like *zi* → *kong* ‘Confucius’ in the context *zi yue* → *kong yue* ‘Confucius says.’⁵ Therefore, in all the other context, *kong yue* renders meaningless, and in any spoken corpora, *kong yue* is acceptable only when it is followed by the very idiomatic phrase.

The system used in this study for classifying types of errors was developed from traditional systems used for defining errors in English. Jaeger (2005) in her corpus made use of the following three dimensions: 1) the linguistic component or stage involved in the error, 2) the directionality of the error, and 3) the form of the error, and similar criteria will be used for the study as well.

The following examples briefly show the figure in the first author’s corpus.

A. Paradigmatic substitutions of two content words (N=660)

- (1) *canguan* → *canjia*
visit participate
‘visit’ → ‘participate’

In this example, the verb ‘participate’ is substituted for ‘visit.’

⁵ Any word occurring in the idiomatic phrase went wrong is considered as an error. In the collocational phrase, the very idiomatic phrase sticking to *kong yue* only in the phrase *kong yue cheng ren, meng yue qu yi* (孔曰成仁孟曰取義); other than this single phrase, we used *zi yue*.

B. Paradigmatic substitutions of two function words (N=70)

- (2) zhe tian → na tian
this day that day
'this day' → 'that day'

This is an example of replacing the target word *zhe* 'this' with 'that', both being the functional category determiner.

C. Paradigmatic substitutions of content-function words (N=28)

- (3) wo fang → wo-men
our side I -Plural
'our side' → 'we/us'

In this example, the pronoun 'we/us' is substituted for the noun 'our side'.

D. Paradigmatic blends of two words (N=54)

- (4) hen man/ben → hen ban
very slow /stupid
'very slow/stupid' → meaningless

In this example, the two lemmas are activated at the same time; they are phonologically blended together and inserted into the same utterance.

E. Syntagmatic blends of two words (including content and function combined) (N=68)

- (5) wo meiyou jiang-dao → wo miu jiang dao
I not.have talk-arrive
'I haven't talked about (the point)...' → meaningless

In this example, the syntagmatic string collapses into a shorter utterance with at least one rhythmic beat (i.e., vowel) omitted.

F. Syntagmatic anticipations of words (N=312)

- (6) ji suo bu yu, wu shi yu ren → ji suo bu shi, wu shi yu ren
self SUO not want not give YU person give give
'Do unto others as you would have them do unto you' → meaningless

In this example, the word 'give' is anticipated and substituted for the word 'want'.

G. Syntagmatic perseverations of words (N=140)

- (7) Babi jiu shi wawa → Babi jiu shi Babi
Barbie just is doll
'Barbie is just a doll' → 'Barbie is just Barbie'

In this example, the word 'Barbie' is perseverated and substituted for the word 'doll'.

H. Syntagmatic reversals of words (N=17)

- (8) Zhen shi duanzhan de kuaile → Zhen shi kuaile de duanzhan
really is short DE happiness happiness short
'(This) is indeed a short period of happiness.' → meaningless

In this case, the adjective 'short' and the noun 'happiness' are reversed.

I. Malapropisms (N=41)

- (9) k^han51 pu51 te^{hi}21 → k^han51 pu51 te^{hi}n55
look not rise look not clear
'despise' → 'can't see clearly'

In this case, typed in IPA, the two words are not semantically related but they share phonological similarities; the two syllables both start with the same consonant and have the same vowel.

J. Environmental errors (N=11)

(10) wo zui xihuan jita → wo zui xihuan niupai
I most like guitar I most like steak

In this case, the two words are nouns, but they are not semantically related. The speaker was talking about the musical instrument, but made the error while he was watching TV, and the named object was physically present in the environment.

In this article, we are interested in the lexical-semantic relationships between the errors and the intended lemma, so lexical substitution errors in a paradigmatic order will be the main data to be examined. Paradigmatic blends (word blend) errors, lexical substitution errors in the syntagmatic order (i.e., errors involving syllable contraction), phonological relationships in lexical substitution errors (i.e., malapropisms), and environmental errors, which will be addressed and provided with examples in the methodology section, will be excluded for the present study, and left for the future research. There are five reasons why paradigmatic-lexical substitution errors are the only relevant data to be looked into in this study. First, most of the afore-mentioned studies do not include errors other than the paradigmatic-lexical substitution type, and even though some studies do, their findings may not present any figures or may not have a similar classification system, so the findings cannot be compared quantitatively with those of other studies. Secondly, word blend errors are excluded since Levelt (1989) stated that the target-error word pairs in word blend errors have conceptual-semantic relationships with each other, and word pairs involving lexical substitution have the lexical-semantic links, and the issue on conceptual-semantic relationships is not the focus of this paper.⁶ Besides, when two lexical items are blended together, the resultant form usually becomes a non-word. Thirdly, syntagmatic lexical errors are excluded because this type of errors occurs when both target and interfering units are in the discourse context. This may occur because it has not been associated with a functional/syntactic slot, or the wrong functional/syntactic frame has been selected (Garrett,

⁶ The distinction between ‘conceptual semantics’ and ‘lexical semantics’ defines a major controversy surrounding the organization of the lexicon.

1982, 1984). This is why the syntagmatic errors are not discussed in most of the relevant studies as mentioned above. Fourthly, malapropisms are excluded because this type of errors involves phonological relationships in lexical substitution errors so the word pairs do not have semantic meanings in common. Besides, comparisons of models regarding whether the words are organized into network with connections between units based on semantic and phonological relatedness are not the focus under study. Finally, environmental errors are also excluded because environmental words result from environmental contamination analyses. Environmental errors occur when the referent of the error word is physically present in the environment. These errors usually involve the common noun names for some objects visible in the environment. In models of speech production, visual representation is always involved in the process and in this paper, we are mainly concerned the formulation stage in the context of psycholinguistic and neurolinguistic models of sentence production.

Paradigmatic substitution errors are by far the most common errors in every study which has been reported (Hotopf, 1983; Jaeger, 2005). There is an overall pattern in the proportion of paradigmatic to syntagmatic errors in the corpora (Jaeger; Hotopf; Meringer). The Jaeger adult figures show roughly 75% paradigmatic and 25% syntagmatic errors, and Hotopf's English data has similar figures. The Meringer's numbers are somewhat different since they have 66% paradigmatic and 34% syntagmatic errors. The similar figure is also made in the Mandarin corpus where there are 60% paradigmatic and 40% syntagmatic errors.

This study aims at addressing those research questions and presenting various patterns in different categories by looking at word-pair errors in paradigmatic order, the type of which is found to be the largest percentage of errors in all the reported speech error corpora cross-linguistically. This study involves semantic relationship of content words and proper names only since content and function words do not behave the same way in speech production (Garrett, 1980).

4. RESULTS AND DISCUSSION

In order to see how closer the word pairs are, the type of semantic relationships are rank-ordered. The following table shows the number of each type of semantic relationship involved in semantically related lexical error in our corpus, organized by a data set.

Table 1. Semantic relationships in lexical errors

| Semantic relationships | Number |
|------------------------|-----------|
| Coordinates | 242 (39%) |
| Co-hyponyms | 166 |
| Co-members of a set | 72 |
| Co-paronyms | 4 |
| Subsumatives | 64 (10%) |
| Subord./superord. | 32 |
| Set/member | 12 |
| Whole/part | 20 |
| Contrastive | 46 (7%) |
| Binary antonyms | 26 |
| Converses | 4 |
| Gradable antonyms | 16 |
| Near-Synonyms | 83 (13%) |
| Association | 181 (29%) |
| Metonymic | 76 |
| Connotation | 4 |
| Causal | 14 |
| Shared features | 87 |
| Total | 616 |

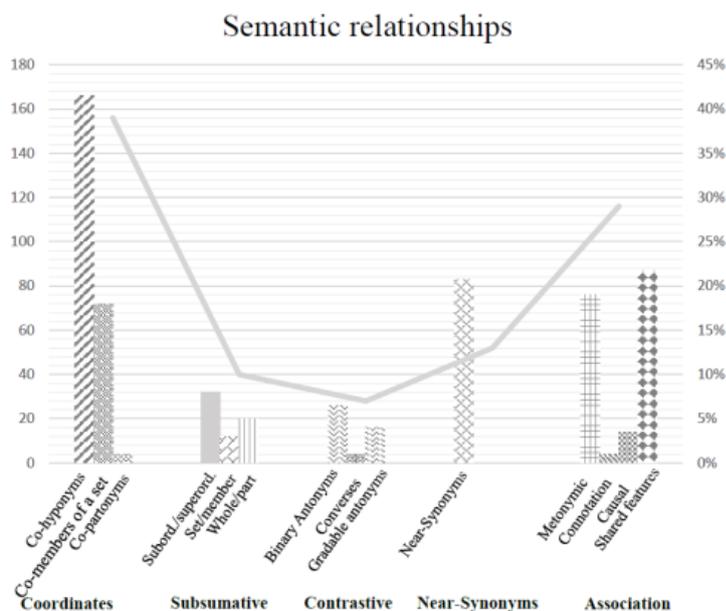


Figure 1. Semantic relationships in categories

The target-error pairs are analyzed in terms of the five types of semantic relationship between the target word and the error word, namely, coordinates, subsumatives, contrastive, near-synonyms, and association. There is a statistically significant difference in these five categories ($\chi^2 (4) = 213.7, p < .01$). Except for near-synonyms, which do not have sub-categories, the rest of the sub-categories all showed statistically significant differences; Coordinates ($\chi^2 (2) = 164.1, p < .01$), Subsumatives ($\chi^2 (2) = 9.5, p < .01$), Contrastive ($\chi^2 (2) = 15.8, p < .01$), and Association ($\chi^2 (3) = 118.6, p < .01$). However, there is no significant difference between subsumatives and near-synonyms ($\chi^2 (1) = 2.5, p > .05$). Examples of each type and subtype of word-pair substitutions from our corpus are provided below.

(11) Coordinates_Co-hyponyms

dongci → mingci
 verb noun
 ‘verbs’ → ‘nouns’

In this example, the target word and the error word are analyzed as co-hyponyms since they share the same superordinate, namely the inventory of the part-of-speech classes of a language, and thus this word pair presents a coordinate relation between the two words.

- (12) Coordinates_Co-members of a set
jiaoshi → bangongshi
classroom office
'classroom' → 'office'

In this example, the target and error words, classroom and office, describe entities which are in the same school area, and are analyzed as co-members of a set. Therefore, the two words have a coordinate relationship.

- (13) Coordinates_Co-paronyms
lubian → lumian
road.side road.surface
'the side of the road' → 'the surface of the road'

In this example, the target word *lubian* 'the side of the road' and the error word *lumian* 'the surface of the road' are parts of the road, and are therefore classified as being co-paronyms, sharing a coordinate relationship.

- (14) Subsumatives_Subordinate/superordinate
xiaohai → ren
child person
'children' → 'human beings'

In this example, the target word 'children' denotes a kind of human beings, which is denoted by the error word. Therefore, this target-error pair is analyzed as showing a subsumative relation.

(15) Subsumatives_Set/member

jiaoshi → xuexiao
classroom school
'classroom' → 'school'

In this example, classroom, denoted by the target word, is a place unit within the school area, denoted by the error word. Therefore, the target word and error words are classified as having a set and member relation, i.e., a subsumative relation.

(16) Subsamatives_Whole/part

dingshuji → dingshuzhen
stapler staples
'stapler' → 'staples'

In this example, staples, denoted by the error word *dingshuzhen* 'staple', are a part of a stapler, denoted by the target word *dingshuji* 'stapler'. Therefore, the target word and the error word are classified as having a whole and part relation, sharing a subsumative relation.

(17) Contrastive_Binary antonyms

jieshou → jujue
accept reject
'accept' → 'reject'

In this example, the action to accept, denoted by the target word *jieshou* 'accept', entails the lack of the action to reject, denoted by the error word *jujue* 'reject'. Therefore, the target word and error word are analyzed as sharing the relation of binary antonyms, a contrastive relation.

(18) Contrastive_Converses

mai → mai
sell buy
'sell' → 'buy'

In this example, the target word *mai* 'sell' and the error word *mai* 'buy' denote essentially the same conceptual relation between two entities. A

classic example of such a relation is illustrated by *learn* vs. *teach* in the methodological section. Therefore, the target and the error words in this example are analyzed as converses, presenting a contrastive relation.

(19) Contrastive_Gradable antonyms

leng → re
 cold hot
 ‘cold’ → ‘hot’

In this example, the notion conveyed by the target word *leng* ‘cold’ and that conveyed by the error word *re* ‘hot’ mark the two ends of a perceived continuum that has intermediate points. The negation of ‘cold’ does not entail the meaning of ‘hot’, which is at the other end of the scale. Therefore, the target word *leng* ‘cold’ and the error word *re* ‘hot’ are analyzed as being gradable antonyms, sharing a contrastive relation.

(20) Near-synonyms

wanxiao → xiaohua
 joke joke
 ‘joke’ → ‘joke’

In this example, both the target word *wanxiao* ‘joke’ and the error word *xiaohua* ‘joke’ have the meaning of jokes. Therefore, they are analyzed as near-synonyms.

(21) Association_Metonymic

xinxiang → youtong
 letterbox post-box
 ‘letterbox’ → ‘post-box’

The target word *xinxiang* ‘letterbox’ refers to a small box on one’s door where one’s letters are delivered, and the error word *youtong* ‘post-box’ refers to a metal box in a public place, where one puts letters to be collected. When discussing the handling of letters, they tend to appear together and are thus analyzed as sharing a metonymic relation, thus presenting an associated relationship.

(22) Association_Connotation

gui → hao
expensive good
'expensive' → 'good'

In this example, being expensive, denoted by the target word *gui* 'expensive', has a strong connotation of being good in quality, denoted by the error word *hao* 'good'. Therefore, the target word *gui* 'expensive' and error word *hao* 'good' are analyzed as being related by connotation, sharing an associated relationship.

(23) Association_Causal

zhaodao → jiuchu
search.arrive rescue.out
'find' → 'rescue'

In this example, the target word *zhaodao* 'find' describes the event of having found something/someone and the *jiuchu* 'rescue' describes the event of rescuing someone. The finding event must take place before the rescuing event. Therefore, the target word *zhaodao* 'find' and error word *jiuchu* 'rescue' are analyzed as referring to concepts that are associated by a link of causality, thus sharing an associative relationship.

(24) Association_Shared features

duanyai → duanceng
cliff fault.line
'cliff' → 'fault line'

In this example, the target word *duanyai* 'cliff of displacement' and the error word *duanceng* 'fault' share the feature of being a geological phenomenon that involves dislocation. Capturing this semantic association relation, the target word and the error word are analyzed as sharing features, presenting an associated relationship.

In the error corpus, when the error word has a high degree of semantic similarity to the target word, it seems to be erroneously selected as the spoken output, suggesting what counts as closer semantic relationship. In

our study, the most frequent type of semantic relationship is coordinates (39%), followed by association (29%), the next two, near-synonyms (13%) and subsumatives (10%), while contrastive (7%) is the least common type. The result suggests that the closest lexically linked words involve the coordinate relationships.

The semantic links between the two lemmas in the lexicon have been supported by word association tasks, priming experiments, other psycholinguistic tasks as well as speech-error studies. In the word association task in Clark's (1970) study, antonymic relationships are considered the closest since they are the most commonly produced association responses, followed by coordinates, then superordinate/subordinate relationships, and finally near-synonyms and associated type relationships. However, this finding is not found in all speech-error studies regarding the lexical activation properties. We compared the results of the study to several other studies of lexical errors. Our results were very similar in many respects to other studies in terms of proportion of types of semantic relationships.

In Mandarin lexical errors, the figure does not confirm the study found in the word association task; in fact, it confirms the data found in other speech error studies where coordinate-type relationships seem to be the most common, followed by associated types, and then near-synonyms and subsumatives. Contrastive relationships are the least common. Jaeger (2005) pointed out that the different ranking between the word-association task and speech error studies might be due to the fact that the former is a metalinguistic task and the subjects are very likely to select a word occurring in an oppositional pair with the target words in a collocational phrase, but in speech error studies, the speaker is activating a set of concepts relevant to the proposition being formulated, and then a set of lemmas which appropriately formulate these concepts.

There are some striking similarities between the proportion of different types of semantic relationships found in the Mandarin corpus and other studies. For the Jaeger and the Hotopf studies, the figures for coordination are 43%-49%. The Harley and MacAndrew figure may be higher (61.5%) because this category combined coordinates, subsumatives and near-synonyms. The next common type is association (14.5%-34.5%) in all studies except Hotopf, who found contrastive type is the next

common in his dataset. The Mandarin error corpus confirms the studies in Harley and MacAndrew (2001) and Jaeger (2005). Errors of near-synonyms and subsumative types are more common than antonyms in Mandarin. Jaeger (2005) suggested the higher rate of coordinates and association in all studies might be due to the fact that the two lexical items involved are at the same level of specificity. When a speaker is planning an utterance and planning to activate the set of lexical items relevant to the concepts in the proposition, this activation is most likely to spread to other lexical items which are related to the target lexical items by sharing a number of semantic features, and sharing values on these features to the same level of specificity in the context. The relationship of co-members of a set is more concrete than that of co-hyponyms. This is because the semantic concept of the 'set' involved in the former semantic relation is more concrete than the concept of the superordinate involved in the latter relation.

As with near-synonyms or the subsumative type, it is difficult for the hearer to know whether a speaker has accidentally substituted a more general for a more specific term. In near-synonyms, even two words might generally refer to the same entity or activity or state of affairs, but they have different selectional restrictions or scope requirements. Errors of this type may sometimes occur and it is not easy for the hearer to detect such an error. In subsumative errors, the word pairs involve either the higher-level term such as superordinate/set/whole or lower-level term such as subordinate/member/part, and the hearer might not notice the existence of the error due to the word pairs having the same reference. This also confirms the hypothesis in Jaeger for an elaborated network of connections in the lexicon in the vertical dimension.

As for the contrastive relationship, a speaker violates the specificity with which the speaker intended to refer to the entity or action involved. The figure in this study is very similar to that reported in Jaeger (2005) and is much lower than that in the rest of the cited studies, but generally speaking, contrastive relationships are the least common type in all the other studies which have previously looked at lexical substitution errors. Hotopf (1980) found the exception in which contrastive errors weight more than other types of errors. This might be due to the different classification system where he included both paradigmatic and

syntagmatic errors and he looked into the content words only. Moreover, the figure also showed that word pairs involving subsumative relationships or near-synonyms did occur, and this is contradictory to what Levelt (1989) claimed for no such data have been found in his dataset.

The most frequent lexical substitution errors are the ones that are semantically close to their targets, suggesting the semantic links among the lemmas in the lexicon. It is true that the target-error pairs can have the following four features: [+phonological, +semantic], [-phonological, +semantic], [+phonological, -semantic], and [-phonological, -semantic]. In the dataset, the most common lexical errors involved the feature [-phonological, +semantic], and the errors sharing the rest of features were not common. Wan and Tang (in progress) classify these four types of lexical errors, and suggest that semantic activation spreads to the target lemma that can have either phonological or semantic correlates, suggesting that lemmas are both stored phonologically and semantically in the mental lexicon.⁷

When lexical substitution errors occur and when a word is semantically substituted for the other without involving any discourse, contextual or collocational proposition, they reflect the failure in the processing of selecting a semantic representation of a word. By looking at the semantic relationships between such word pairs, it will be possible to see how semantic relatedness is represented in lexical selection processes. The data are consistent with the description of the formulation stage in the context of psycholinguistic and neurolinguistic models of sentence production (Garrett, 1975, 1982, 1984; Levelt, 1989; Dell 1986, 1988). The most likely explanation is that the semantic links between the two lemmas in the lexicon have caused the error. The activation spread from the target lemma to the erroneous lemma which shares semantic properties, and is in the various types of relationships with the target lemma.

⁷ We thank the anonymous reviewer who pointed out the importance of lexical errors with phonological relatedness.

5. CONCLUSION

The present study looked at various patterns involving paradigmatic lexical errors in Mandarin by presenting how Mandarin speakers produce the lexical items in relation to erroneous selection and how the erroneously produced words are in place of the target. A detailed taxonomy of lexical-semantic relationship between the target and error words has been adopted. We have shown when two lexical items are substituted for each other, the word pairs who have semantic relatedness are more likely to interact with each other, and semantically related associates are very likely to be involved in the word pairs.

The findings show that lexical substitution errors in paradigmatic fashion involve close semantic relationships, and coordinate and associated features are very common; the number confirms the studies Harley and MacAndrew (2001) and Jaeger (2005); however, errors of near-synonyms and subsumative types are more common than antonyms, a finding which is similar to that of Jaeger (2005). This finding supports the claim that when two lexical items are substituted for each other, they are related by sharing a number of semantic features and sharing values on these features to the same level of specificity in the context.

Although this study does not entail any estimate of the degree of relatedness that occurs between target and word pairs, it distinguishes among types of semantic relations and shows the taxonomy of semantic relationships between the word pairs. In the past, a psycholinguistic experiment such as word priming or association tasks generally investigated words via associative links of varying strengths in the mental network by looking at the effects of semantic relatedness of words. Although this study does not present a model of speech production by including all types of speech errors, linking the interaction among phonology, morphology, phrasal units and syntactic constituents, or even taking frequency effect for consideration, it shows between-lemma links in the mental lexicon of Mandarin speakers, supporting the separation of concept and lemmas in the representation and processing. In the future studies, Wan and Tang (in progress) are going to show the semantic connections between lemmas in Mandarin lexicon for eliminating the artifact of word frequency effect in free word association tasks. Moreover,

the series of this work will be able to focus on a larger picture regarding syntagmatic lexical errors, which are more related to discourse and contextual proposition, word blends in both paradigmatic and syntagmatic orders, which occur in conceptual lexicalization, and malapropisms, which shows the phonological relationships in word pairs that do not involve any semantic relationships.

REFERENCES

- Ahrens, K. 1994. Classifier production in normals and aphasics. *Journal of Chinese Linguistics*, 22(2), 202-247.
- Ahrens, K. 2012. Finding what you expect to see: Theoretical modeling in psycholinguistics. *Journal of Social Sciences*, 8(3), 372-380.
- Alario, F. X. and Cohen, L. 2004. Closed-class words in sentence production: Evidence from a modality-specific dissociation. *Cognitive Neuropsychology*, 21(8), 787-819.
- Arnaud, P. J. 1999. Target-error resemblance in French word substitution speech errors and the mental lexicon. *Applied Psycholinguistics*, 20(02), 269-287.
- Berg, T. 1987. *A Cross-linguistic Comparison of Slips of the Tongue*. Bloomington, IN: Indiana University Press.
- Bock, J. K. and Levelt, W. J. M. 1994. Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics*, pp.945-984, Orlando, FL: Academic Press.
- Chang, Y., Lin, C. J. C. and Ahrens, K. 2015. Conventionalization of lexical meanings and the role of metaphoricity: Processing of metaphorical polysemy using a cross-modal lexical priming task. *Language and Linguistics*, 16(4), 587-614.
- Clark, H. 1970. Word associations and linguistic theory. In J. Lyons (Ed.), *New Horizons in Linguistics*, pp.271-286, Middlesex, England: Penguin Books.
- Coltheart, M. 1980. Iconic memory and visible persistence. *Perception and Psychophysics*, 27, 240-247.
- Cutler, A. 1982. The reliability of speech error data. In A. Cutler (Ed.), *Slips of the Tongue and Language Production*, pp.7-28, Amsterdam: Mouton.
- Dell, G. 1984. Representation of serial order in speech: Evidence from the repeated phoneme effect in speech errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 222-233.
- Dell, G. S. 1986. A spreading activation theory of retrieval in sentence production. *Psychological Review*, 93, 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(2), 124-142.
- Dell, G. and Reich, P. 1980. Toward a unified theory of slips of the tongue. In V. Fromkin (Ed.) *Errors in Linguistic Performance: Slips of the Tongue, Ear, Pen, and Hand*, pp.273-286, NY: Academic Press.
- Dell, G. and Reich, P. 1981. Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior*, 20(6), 611-629.
- Dell, G., Juliano, C. and Govindjee, A. 1993. Structure and content in language production: A theory of frame constraints in phonological speech errors. *Cognitive Science*, 17, 149-195.
- Fromkin, V. A. 1973. *Speech Errors as Linguistic Evidence*. The Hague: Mouton.
- Fromkin, V. A. 1980. *Errors in Linguistic Performance: Slips of the Tongue, Ear, Pen, and Hand*. NY: Academic Press.

- Fromkin, V. A. 1983. Aspects of a model of speech production: Evidence from speech errors. In van den Broecke, M., van Heuven, M. and Zonneveld, W. (Eds.), *Sound structures: studies for Antonie Cohen*, pp.105-112, Dordrecht: Foris.
- Fromkin, V. A. 1988. Grammatical aspects of speech errors. In F. Newmeyer (Ed.), *Linguistics: The Cambridge Survey, V. II: Linguistic Theory: Extensions and Implications*, pp.117-138, Cambridge, UK: Cambridge University Press.
- Garrett, M. F. 1975. The analysis of sentence production. In G. Bower (Ed.), *Psychology of Learning and Motivation*, Vol. 9, pp.133-177, NY: Academic Press.
- Garrett, M. F. 1980. Levels of processing in sentence production. In B. Butterworth (Ed.), *Language Production*, Vol. 1, pp.177-220, London: Academic Press.
- Garrett, M. F. 1982. Production of speech: Observations from normal and pathological language use. In A. Ellis (Ed.), *Normality and Pathology in Cognitive Functions*, pp.19-76, London: Academic Press.
- Garrett, M. 1984. The organization of processing structure for language production. In D. Caplan, A.R. Lecourse, and A. Smith (Eds.), *Biological Perspectives on Language*, pp.172-193, Cambridge: MIT Press.
- Garrett, M. 1993. Errors and their relevance for models of language production. In G. Blanken, J. Dittmann, H. Grimm, J. Marshall and C. Wallesch (Eds.), *Linguistic Disorders and Pathologies*, pp.72-92, Berlin: de Gruyter.
- Garnham, A., Shillcock, R., Brown, G., Mill, A. and Cutler, A. 1982. Slips of the tongue in the London-Lund corpus of spontaneous conversation. In A. Cutler (Ed.), *Slip of the Tongue and Language Production.*, pp.251-263, Berlin etc. : Mouton Publishers.
- Gong, S. P., Ahrens, K. and Huang, C. R. 2008. Chinese word sketch and mapping principles: A corpus-based study of conceptual metaphors using the BUILDING source domain. *International Journal of Computer Processing of Languages*, 21(1), 3-17.
- Gruber, J. 1976. *Lexical Structure in Syntax and Semantics*, North Holland, New York.
- Harley, T. and MacAndrew, S. 2001. Constraints upon word substitution speech errors. *Journal of Psycholinguistic Research*, 30, 395-418.
- Hotopf, W. 1980. Semantic similarity as a factor in whole-word slips of the tongue. In V. Fromkin (Ed.), *Errors in Linguistic Performance: Slips of the Tongue, Ear, Pen, and Hand*, pp.97-109, NY: Academic Press.
- Hotopf, W. 1983. Lexical slips of the pen and tongue: What they tell us about language production. *Language Production*, 2, 147-199.
- Huang, C. R., Ahrens, K. and Chen, K. J. 1998. A data-driven approach to the mental lexicon: Two studies on Chinese corpus linguistics. *Bulletin of the Institute of History and Philology*, 69(1), 151-179.
- Huang, C. R., Ahrens, K., Chang, L. L., Chen, K. J., Liu, M. C. and Tsai, M. C. 2000. The module-attribute representation of verbal semantics: From semantics to argument structure. *International Journal of Computational Linguistics and Chinese Language Processing*, 5(1), 19-46.

- Huang, Chu-Ren, Tseng, Elanna I. J., Tsai, Dylan B. S. and Murphy, B. 2003. Cross-lingual portability of semantic relations: Bootstrapping Chinese WordNet with English WordNet relations. *Language and Linguistics*, 4(3), 509-532.
- Jaeger, J. J. and Wilkins, D. 1993. Some Lexical and Semantic Properties of Young Children's Slips of the Tongue. Unpublished manuscript. SUNY-Buffalo.
- Jaeger, J. J. and Wilkins, D. 2005. Semantic relationships in lexical errors. In Jaeger J. J. (Ed.), *Kid's Slips: What Young Children's Slips of the Tongue Reveal about Language Development*, pp.311-384, Mahwah, NJ: Lawrence Erlbaum Associates.
- Jaeger, J. J. 2005. *Kids' Slips: What Young Children's Slips of the Tongue Reveal about Language Development*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Kess, J. F. 1992. *Psycholinguistics: Psychology, Linguistics, and the Study of Natural Language*. Amsterdam: John Benjamins Publishing.
- Levelt, W. J. 1983. Monitoring and self-repair in speech. *Cognition*, 14(1), 41-104.
- Levelt, W. J. 1989. *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.
- Levelt, W. J. 1992. Accessing words in speech production: Stages, processes, representations. *Cognition*, 42, 1-22.
- Levelt, W., Roelofs, A. and Meyer, A. 1999. A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(1), 1-38.
- Liu, M. C. and Chiang, T. Y. 2008. The construction of Mandarin VerbNet: A frame-based study of statement verbs. *Language and Linguistics*, 9(2), 239-270.
- Liu, M. C., Chiang, T. Y. and Chou, M. H. 2005. A Frame-based Approach to Polysemous Near-synonymy: the Case with Mandarin Verbs of Expression. *Journal of Chinese Language and Computing*, 15(3), 137-148.
- Liu, M. and Hu, C. 2008. Conceptual schema as semantic link: A frame-based study of Mandarin cognition verbs. *International Journal of Computer Processing of Languages*, 21(1), 55-75.
- Martin, N., Dell, G. S., Saffran, E. M. and Schwartz, M. F. 1994. Origins of paraphasias in deep dysphasia: Testing the consequences of a decay impairment to an interactive spreading activation model of lexical retrieval. *Brain and Language*, 47(4), 609-660.
- Myers, J. and Gong, S. 2002. Cross-morphemic predictability and the lexical access of compounds in Mandarin Chinese. *Folia Linguistica*, 36(1-2), 65-96.
- Nooteboom, S. G. 1973. The tongue slips into patterns. In V. Fromkin (Ed.), *Speech Errors as Linguistic Evidence*, pp.87-95, The Hague: Mouton.
- Packard, J. L. 2000. *The Morphology of Chinese: A Linguistic and Cognitive Approach*. Cambridge: Cambridge University Press.
- Seto, K. I. 1999. Distinguishing metonymy from synecdoche. *Metonymy in Language and Thought*, 4, 91-120.
- Shattuck-Hufnagel, S. 1979. Speech errors as evidence for a serial ordering mechanism in sentence production. In W. E. Cooper and E. C. T. Walker (Eds.), *Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett*, pp.295-342, Hillsdale, N. J.: Erlbaum.

- Stemberger, J. 1983. *Speech Errors and Theoretical Phonology: A review*. Unpublished manuscript, Carnegie-Mellon University. Distributed by the Indiana University Linguistics Club, Bloomington, IN.
- Stemberger, J. 1985. The reliability and replicability of naturalistic speech error data: A comparison with experimentally induced errors, *Research on Speech Perception Progress Report, 11*, 235-261.
- Stemberger, J. 1989. Speech errors in early child language production. *Journal of Memory and Language, 28*, 164-188.
- Stemberger, J. 1993. Spontaneous and evoked slips of the tongue. In G. Blanken, J. Dittmann, H. Grimm, J. Marshall and C. Wallesch (Eds.), *Linguistic Disorders and Pathologies*, pp.53-65, Berlin: de Gruyter.
- Tsai, M. C., Huang, C. R., Chen, K. J. and Ahrens, K. 1998. Towards a representation of verbal semantics: An approach based on near synonyms. *Computational Linguistics and Chinese Language Processing, 3*(1), 61-74.
- Wan, I. P. 1999. *Mandarin Phonology: Evidence from Speech Errors*. Ph.D. dissertation, SUNY-Buffalo.
- Wan, I. P. 2007a. Mandarin speech errors into phonological patterns. *Journal of Chinese Linguistics, 35*(1), 185-220.
- Wan, I. P. 2007b. *A Phonological Investigation in Speech Errors and Aphasic Speech in Mandarin*, Taipei: Crane Publishing.
- Wan, I. P. and Tang, M. (In progress) A corpus study of lexical speech errors in Mandarin. Manuscript.
- Wijnen, F. 1992. Incidental word and sound errors in young speakers. *Journal of Memory and Language, 31*, 734-755.
- Yong, M. Y. 2015. *A Comparison of Mandarin Near Synonyms sheng1 ming4 (生命) and sheng1 huo2 (生活) with English 'life': A Contrastive Study Based on Corpus and Questionnaire* (Unpublished Master's thesis). National Chengchi University, Taipei, Taiwan.

[Received 2 March 2017; revised 13 October 2017; accepted 5 December 2017]

I-Ping Wan

Graduate Institute of Linguistics and Research Center for Mind, Brain & Learning
National Chengchi University
Taipei, Taiwan 116, ROC
ipwan@nccu.edu.tw

Jen Ting

Department of English
National Taiwan Normal University
Taipei, Taiwan 106, ROC
ting@ntnu.edu.tw

台灣華語語誤之語意關連性

萬依萍¹、丁仁²

¹國立政治大學

²國立臺灣師範大學

本篇研究主要是從語誤語料庫來探討中文詞彙置換的語意關係。從中文詞彙語誤的語意代換中可以大致上證實英語及德語的現象，由此可見跨語言間相似的情形。這其中的相似之處在於詞彙置換語誤通常是與語意的關連度有關。目標與語誤詞彙(target-error words)之間會有相似的語意特徵或者呈現語意學裡分類架構中的關聯性。從五種分類架構中來看，語意並列(coordinate)是最常見的情況，再來就是關聯性(association)，對比關係(contrastive)則是最少見的。這樣的研究結果支持了對於兩個詞彙互相置換時，彼此都有一定程度上相似的語意特徵或是共同擁有在情境上的某些特質的證論。中文裡的詞彙錯誤總體來說與國外一些跨語言研究中所得到的結論大致吻合。也就是說，詞彙裡的兩個詞目(lemma)之間的語意關聯即是構成置換錯誤的主因。從中文語誤語料的分布類別可以清楚的看出心理語言學及神經語言學在語言產製時直接及透明的歷程，與詞彙選擇呈現獨立進行的過程。

關鍵字：詞彙置換語誤、語意關連性、語誤、台灣華語