



Universiteit
Leiden
The Netherlands

A psycholinguistic investigation of speech production in Mandarin Chinese

Wang, M.

Citation

Wang, M. (2017, July 5). *A psycholinguistic investigation of speech production in Mandarin Chinese*. Retrieved from <https://hdl.handle.net/1887/51344>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/51344>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/51344> holds various files of this Leiden University dissertation

Author: Wang, M.

Title: A psycholinguistic investigation of speech production in Mandarin Chinese

Issue Date: 2017-07-05

Chapter 1

Introduction

Speaking a language is a unique capability of human beings. Words, together with their semantic, syntactic and phonological properties, are stored in our mental lexicon (Aitchison, 2012). When we speak, we access the mental lexicon at an amazingly high speed to select the to-be-produced words and to express the meaning in their appropriate phonological forms within the syntactic constraints (Van Turenout, Hagoort, & Brown, 1998). Several influential models have been proposed to capture the underlying mechanisms of language production, in particular speech production. However, these models have mostly drawn evidence from West Germanic languages. In recent decades, studies researching the speech production of languages with a logographic script have questioned the accountability of current speech production models. For instance, while orthographic vs. phonological forms are less differentiated in West Germanic languages, pure orthographic relatedness has been reported to affect speech production in Mandarin Chinese (Bi, Xu, & Caramazza, 2009; Zhang, Chen, Weekes, & Yang, 2009; Zhang & Weekes, 2009; Zhao, La Heij, & Schiller, 2012). In speech production in Mandarin Chinese, there is also mixed evidence supporting either a syllabic unit of phonological encoding (Chen et al., 2002; O'Seaghdha, Chen, & Chen, 2010) or a sub-syllabic encoding (e.g., Qu, Damian, & Kazanina, 2012; Verdonschot, Lai, Chen, Tamaoka, & Schiller, 2015). This dissertation aimed to bring new insights into these debates by providing behavioral and electrophysiological evidence from speech production in Mandarin Chinese.

In this chapter, first, I will introduce the psycholinguistic models of speech production. Then, I will talk about where the accountability of these models has been questioned and how the dissertation contributes to the understanding of current speech production models.

1.1 A brief introduction to current psycholinguistic models of speech production

In psycholinguistics, the speech production mechanisms are mainly investigated by speech error and picture naming research (see Levelt, 1999 for a review). Although models of speech production differ in the terminology and details about the processing stages, they generally recognize several major processing stages: conceptualization, lemma retrieval, word-form encoding and articulation (e.g., Caramazza, 1997; Levelt, 1992, 1993; Dell & O'Seaghdha, 1991, 1992; the WEAVER++ model, Levelt, Roelofs, & Meyer, 1999a, b; Roelofs, 1992; Roelofs & Meyer, 1998; see Figure 1.1).

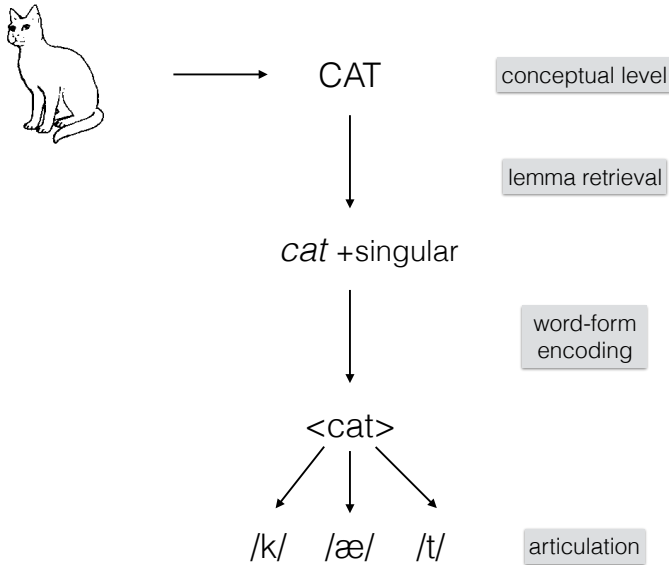


Figure 1.1 Stages of lexical access in WEAVER++ (adapted from Roelofs, 2000).

For instance, when one is asked to name a picture (e.g. a cat), the correct perception of the picture will activate its corresponding concept (e.g. CAT). Notably, the semantically related concepts may be activated as well (e.g., DOG, ANIMAL; Collins & Loftus, 1975; Levelt et al., 1999; Glaser & Dungelhoff, 1984; Indefrey & Levelt, 2004; Starreveld & La Heij, 1995). An alternative possibility is that the perceived concept (e.g. CAT) activates the related semantic features (e.g., FUR, PAW, ANIMAL) in a relatively decomposable manner (Dell & Seaghdha, 1991). The outcomes of the two ways of activation are similar. That is, the semantically related concepts (e.g. DOG) will be activated either directly by the perceived concept (e.g. CAT) or by the activated overlapping semantic features (e.g., FUR, ANIMAL). This conceptual preparation process normally takes up to about 200 ms, according to the

comprehensive meta-analyses of imaging experiments on language production (Indefrey & Levelt, 2004; Indefrey, 2011; see Figure 1.2).

Subsequently, the activated concept (e.g. CAT) will activate its lexical-syntactic representation, i.e. lemma (e.g. cat; the WEAVER++ model, Levelt et al., 1999a, b; Roelofs, 1992; Roelofs & Meyer, 1998). Lemma nodes in the lexical network contains the intrinsic syntactic properties such as grammatical gender, word category etc. The extrinsic properties such as number are activated via the lemma or/and the concept MULTIPLE (see Nickels, Biedermann, Fieder, & Schiller, 2015 for the framework of the lexical-syntactic representation of number). The lemma retrieval process continues till about 275 ms after picture presentation (Indefrey & Levelt, 2004; Indefrey, 2011; see Figure 1.2) and takes place in left middle temporal gyrus (MTG; Schuhmann, Schiller, Goebel, & Sack, 2012). Under certain circumstances, the latency may increase if the semantically related lemma nodes (e.g. dog) are highly activated and compete for lexical selection (WEAVER++; Levelt et al., 1999a, b; Roelofs, 1992; Roelofs & Meyer, 1998; but see e.g. Dell, Schwartz, Martin, Saffran, & Gagnon, 1997 for a non-competitive account), or decrease if the target lemma (e.g. cat) has previously been activated due to repetition priming (e.g., Mitchell & Brown, 1988; Wheeldon & Monsell, 1992).

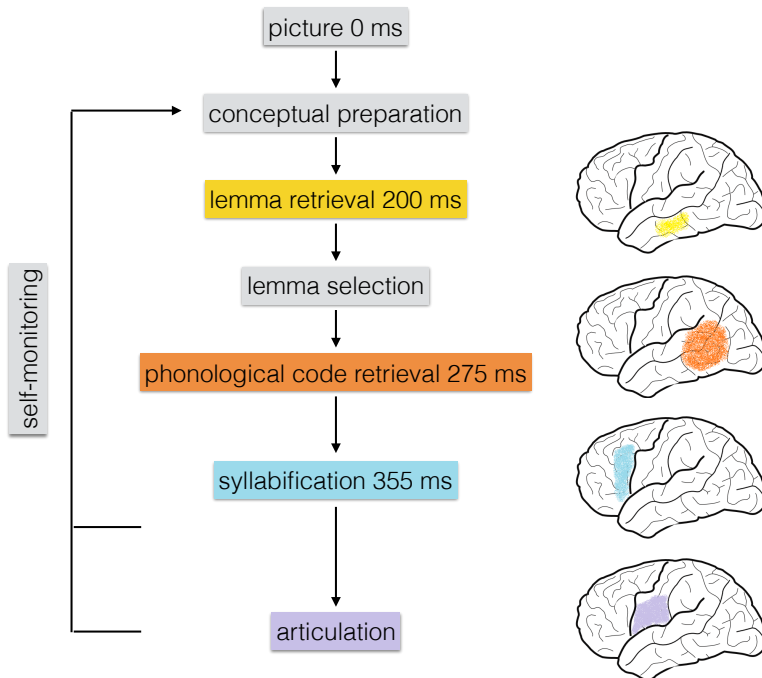


Figure 1.2 Schematic representation of the activation time course of brain areas involved in word production (adapted from Indefrey, 2011).

Following the lemma retrieval stage, the activations flow to the phonological form encoding stage, including phonological code retrieval, syllabification and phonetic encoding (the WEAVER++ model, Levelt et al., 1999a, b; Roelofs, 1992; Roelofs & Meyer, 1998). In West Germanic languages, it is commonly assumed that the phonological segments and metrical frames are activated in parallel and then encoded serially for articulation. This final stage of speech production usually lasts until about 600 ms after the picture presentation (see, Indefrey & Levelt, 2004; Levelt, 2011 for a detailed estimation of specific sub-stages of phonological form encoding; Figure 1.2) and takes place in Broca’s area (Schuhmann, Schiller, Goebel, & Sack, 2009).

1.2 When models based on West Germanic languages meet Mandarin Chinese

It is worth noting that the most influential models of speech production have mainly drawn on evidence from West Germanic languages and orthographic vs. phonological forms are less differentiated (but see Roelofs, 2015 for the modeling of phonological encoding in Mandarin and Japanese spoken word production as well as Mandarin and Japanese versions of WEAVER++). Even within languages with an alphabetic writing system, language systems vary in terms of the depths of orthography (Katz & Frost, 1992). The mechanisms of word-form processing may hence differ across languages and should be accounted for by models of speech production. As the example shown in Figure 1.3, some languages like Macedonian have a shallow orthography, i.e. grapheme and phoneme have a strict one-to-one correspondence. Some other languages like English have a deep orthography, i.e. the degree of consistency and completeness between grapheme and phoneme is much lower (see, e.g. Katz & Frost, 1992). For instance, the rhyme *ear* in the words *bear* and *year* has different pronunciations, i.e. [eə^r] and [ɪə^r], respectively. In languages with a logographic writing system, however, grapheme and phoneme have a highly arbitrary correspondence. Take Mandarin Chinese as an example, the basic unit of the writing system is a character (e.g. 书, ‘book’), and one character usually corresponds to a syllable (e.g. shu1, ‘book’). The number of possible syllables in Mandarin Chinese is limited: about 1,300 syllables including lexical tones (Duanmu, 2002). As a result, there are a large number of homophones, especially at the syllabic/morphonemic level. Brain imaging research has shown that there is a high interactivity of orthography and phonology during homophone judgement (Siok, Perfetti, Jin, & Tan, 2004). Therefore, orthography plays a crucial role in distinguishing homophones and may be involved in speech production in Mandarin Chinese.

<i>Macedonian</i>	<i>English</i>	<i>Chinese</i>
книга	book	书
↓ ↓ ↓ ↓ ↓	↓ ↓ ↓	↓
[knigə]	[buk]	[ʃu]

Figure 1.3 An illustration of the difference in depth of orthography in three exemplar languages.

In language comprehension, it has been found that when Chinese-English bilinguals perceive the presented English word pairs (e.g. train - ham; apple - desk), there is an ERP effect between pairs whose Chinese equivalents are orthographically similar (e.g. 火车 - 火腿) and those that are unsimilar (e.g. 苹果 - 桌子; Thierry & Wu, 2007). Although the study was carried out to investigate the activation of native language during second language comprehension, the results indicate the possibility that the orthographic representation of Chinese words (i.e. Chinese characters) may be activated even when the information is irrelevant for the linguistic tasks that the participants are instructed to perform. In a different line of research, it has been found that a presented character that is orthographically similar (e.g. 庆, qing4, ‘celebration’) facilitates the naming the target picture (e.g. 床, chuang2, ‘bed’; Bi et al., 2009; Zhang et al., 2009; Zhang & Weekes, 2009; Zhao et al., 2012). Nevertheless, there is a debate on when and how orthographic relatedness affects speech production, i.e. facilitating at the word-form encoding stage (Zhao et al., 2012) or facilitating lemma retrieval via an earlier lexical-semantic pathway (Zhang & Weekes, 2009).

In addition, the neural correlates of speech production have been investigated mainly using West Germanic languages with brain imaging measurements (see Ganushchak, Christoffels, & Schiller, 2011 for a review), whereas it is less clear about the underlying neuropsychological mechanisms of speech production in a language with a logographic script like Mandarin Chinese.

Investigations of speech production of Mandarin Chinese contribute to the understanding of current psycholinguistic models of speech production. On the one hand, while the confounds between orthography and phonology make it difficult to interpret the experimental observations (e.g. to separate the contributions of spelling or sound to speech production in languages with an alphabetic script), thanks to the opaque mapping between orthography and phonology, the separate roles of orthography and phonology can be easily addressed in languages with a logographic script. On the other hand, the behavioral and electrophysiological evidence contributes to the understanding of the neuropsychological mechanisms of speech production of languages with a logographic writing system.

This dissertation investigates the specific stages involved in speech production and tests to what extent the current psycholinguistic models of speech production can account for cross-linguistic differences. For instance, in the case of Mandarin Chinese, does orthography contribute to speech production? If so, when and how can orthography affect speech production? Does orthography interact with semantics or phonology in speech production? What are the neural correlates of semantic and phonological processing during speech production in Mandarin Chinese? Are lexical-syntactic features automatically activated in speech production?

1.3 Types of Mandarin Chinese characters

Before introducing the methodology of the experimental research, I will first introduce the major types of Mandarin Chinese characters - simplex and complex characters. Complex characters in this dissertation refer to those that are composed of a semantic radical and a phonetic radical. This kind of character takes up to 80% of the Mandarin Chinese characters (Zhou, 1978; Zhou, Peng, Zheng, Su, & Wang, 2013). For instance, the content word 锤 (chui2, 'hammer') is composed of two radicals. One is the radical on the left: 钅 is called the semantic radical of the character. It is a common semantic radical that usually indicates the character is semantically related to metal. The other radical on the right, i.e. 垂 (chui2, 'suspend'), is called the phonetic radical of the character. The phonetic radical usually indicates the sound of the whole character. A simplex character refers to those that are composed of a single, non-decomposable component (pictographic or ideographic characters), such as 垂 (chui2, 'suspend'). Nevertheless, the indications of semantic and phonetic radicals may not always be as transparent as the given example.

These characteristics make Chinese characters an interesting test case for the possible role of orthography in speech production. Using simplex characters can easily dissociate orthography from semantics and phonology while using complex characters allows us to test possible interactions between orthography and semantics and phonology.

1.4 Experimental paradigms and measurements used in this dissertation

Picture naming has been widely used to investigate speech production. To answer these questions, this dissertation makes use of two picture-naming paradigms that are commonly employed in the field of speech production research. Previous research has demonstrated that orthography affects speech

production but mostly in reading or character naming tasks in languages with an alphabetic script (e.g. Dutch; Roelofs, 2006) as well as languages with a logographic script (e.g., Chinese; Bi, Wei, Janssen, & Han, 2009; Japanese; Yoshihara, Nakayama, Verdonschot, & Hino, 2017). Compared to reading or character naming tasks that rely heavily on the grapheme-to-phoneme transformation, picture naming paradigms capture a more conceptually-driven cognitive process of speech production given the required lexicalization of the concept before phonological encoding (see e.g. Glaser, 1992 for a review of picture naming models and discussions over comparing reading and picture naming). The question of interest is: Without the compulsory grapheme-to-phoneme transformation, can orthography influence the conceptually-driven speech production process?

One of the two paradigms used in this dissertation is the picture-word interference paradigm (e.g., Lupker, 1979; Rosinski, Golinkoff, & Kukish, 1975). In this paradigm, participants are asked to name pictures (black-and-white line drawings) while ignoring a distractor word on the picture. By manipulating the relatedness between the distractor word and the target, we observe differences in naming latencies. It has been generally reported that when the distractor (猫, mao1, ‘cat’) and the target (狗, gou3, ‘dog’) belong to the same semantic category, the naming latencies are longer relative to an unrelated condition (窗, chuang1, ‘window’). This is called the semantic interference effect. When the distractor (猫, mao1, ‘cat’) is phonologically related to the target (帽, mao4, ‘hat’), the naming latencies are shorter, relative to an unrelated condition. This is called the phonological facilitation effect (e.g., Glaser & Döngelhoff, 1984; Schriefers, Meyer, & Levelt, 1990; Starreveld, 2000; Starreveld & La Heij, 1995, 1996; see Glaser, 1992; MacLeod, 1991 for reviews of the paradigm). The semantic interference effect and the phonological facilitation effect have been reported in languages with an alphabetic script as

well as languages with a logographic script (see, Bi et al., 2009; Zhao et al., 2012 for the independent orthographic and phonological facilitation effects in Mandarin Chinese; Wong & Chen, 2008, 2009 for the phonological facilitation effect in Cantonese spoken word production; Zhang et al., 2009; Zhang & Weekes for the semantic interference effect as well as the orthographic and phonological facilitation effects in Mandarin Chinese).

The other paradigm is the blocked-cyclic naming paradigm (Damian, Vigliocco, & Levelt, 2001; Belke, Meyer, & Damian, 2005). In this paradigm, target pictures are grouped into homogeneous or heterogeneous blocks. In the homogeneous block, pictures either belong to the same semantic category (e.g., *apple, peach, pear, orange*) or they are phonologically related (e.g., *coat, cat, cook, court*). In the heterogeneous block, pictures are semantically and phonologically unrelated. It has been reported that the naming latencies are longer in the semantically homogeneous blocks than the heterogeneous blocks (e.g., Belke et al., 2005; Damian et al., 2001; Damian & Als, 2005; but see Navarrete, Del Prado, Peressotti, & Mahon, 2014). This is referred to as the semantic blocking effect. Moreover, the naming latencies are shorter in the phonologically homogeneous blocks than the heterogeneous blocks (Damian, 2003; Damian, & Stadthagen-Gonzalez, 2009; but see Damian & Dumay, 2009). This is referred to as the phonological facilitation effect.

It has been noted that “an overt response reflects the output of a large number of individual cognitive processes, and variations in reaction time (RT) and accuracy are difficult to attribute to variations in a specific cognitive process. ERPs, in contrast, provide a continuous measure of processing between a stimulus and a response, making it possible to determine which stage or stages of processing are affected by a specific experimental manipulation.” (Luck, 2005, p. 21). Event-related potential (ERP) experiments have been carried out extensively in linguistic research. However, the majority of the

experiments investigate language perception processes and covert language production. This is mainly due to the concerns about muscle movements involved in language production which can distort the ERP signals and consequently make the acquired data unreliable. However, an increasing number of recent studies have investigated the functional characters of speech production with electrophysiological measurements and shown that artifact-free ERP signals can be measured up to 400 ms post-stimulus presentation (Ganushchak et al., 2011). The reliability of electrophysiological measurement with overt speech production calls for more research to provide fine-grained data with high temporal resolution to reveal the underlying mechanisms of speech production.

In this dissertation, we not only measured the participants naming latencies (i.e. behavioral data) but also their electrophysiological activities (i.e. EEG data) so as to provide more insights to understanding the speech production mechanisms as well as the inherent components of the experimental paradigms.

1.5 Overview of the experimental chapters

In general, Chapters 2 and 3 focus on the orthographic effect on speech production in Mandarin Chinese and Chapters 4 and 5 focus on the neural correlates of speech production in Mandarin Chinese.

Chapter 2 tests whether orthography contributes to speech production in Mandarin Chinese. Specifically, we asked participants to name pictures of simple objects while presenting Chinese characters very briefly (75 ms) before the pictures. We observed that orthographically related characters facilitated the picture naming process, i.e. shorter naming latencies.

Chapter 3 focuses on a more specific debate that whether orthography can affect speech production at an early stage via the lexical-semantic pathway. We firstly used the complex characters to test possible interactions between orthography and semantics and then simplex characters to re-capture the time course of semantic, phonological and orthographic processing in speech production. We observed that orthography affected speech production at a similar stage to phonology, subsequent to semantic processing.

Chapter 4 investigates the neural correlates of semantic and phonological processing in Mandarin Chinese speech production. We observed that the semantic factor started to affect electrophysiological activities from 200 ms and phonological factor from 350 ms. We also observed correlations between the behavioral effects and the electrophysiological effects. Phonological facilitation was also observed with sub-syllabic overlap, which contributes to the debate concerning the encoding unit of phonological forms during speech production of Mandarin Chinese.

Chapter 5 tests whether the lexical-syntactic features are activated and selected in speech production. Using both behavioral and electrophysiological measurements, we were able to show that the lexical-syntactic feature in question, i.e. the Chinese classifier, was activated but not selected in bare noun speech production of Mandarin Chinese.