A psycholinguistic investigation of speech production in Mandarin Chinese

Wang, M.

A PSYCHOLINGUISTIC INVESTIGATION OF SPEECH PRODUCTION IN MANDARIN CHINESE

Man Wang

A Psycholinguistic Investigation of Speech Production in Mandarin Chinese

Cover design: Ziqing Xu

Print: Ridderprint BV, Ridderkerk

ISBN: 978-94-6299-629-8

A Psycholinguistic Investigation of Speech Production in Mandarin Chinese

PROEFSCHRIFT

ter verkrijging van

de graad van Doctor aan de Universiteit Leiden,

op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,

volgens besluit van het College voor Promoties

te verdedigen op woensdag 5 juli 2017

klokke 11.15 uur

door

Man Wang

geboren te Shandong

in 1987

Promoter:	Prof. dr. N.O. Schiller
Co-Promoter:	dr. Y. Chen
Promotiecommissie:	Prof. dr. J. Doetjes
	Prof. dr. A. Roelofs (Radboud Universiteit)
	Prof. dr. R. Sybesma
	Dr. R. Verdonschot (Waseda Institute for Advanced Study)

Acknowledgement

Writing a PhD dissertation is so much more challenging but rewarding than I had expected. I want to express my deepest gratitude to my supervisors Niels Schiller and Yiya Chen. Thank you for introducing me to such a fascinating subject and I benefit so much from your sharp minds and scopes of knowledge. I want to thank you especially for your constant support and encouragement along this journey.

Special thanks go to my collaborators of my PhD project. Prof. dr. Minghu Jiang, for offering me the 'unlimited membership' of the Brain and Cognitive Neuroscience Lab in Tsinghua University (Beijing) and by all means support for carrying out my field research in Beijing. Dr. Zeshu Shao, I feel so fortunate to have met you and been able to explore our common interests in such an efficient and constructive way. Prof. dr. Antje Meyer, thank you for sharing your knowledge and for your honest and helpful comments.

I am grateful to all my participants in Leiden and China. Thank you for expressing your interests in my experiments and your endurance for repeatedly naming mundane objects and sitting in a small room with an EEG cap trying not to move or blink.

I would not have been able to finish my PhD dissertation with sanity without my colleagues inside and outside LUCL. Bobby, I feel so lucky to have you as my colleague and friend while carrying out my EEG experiments in the Pieter de la Court building and the Lipsius building. Thank you for translating my Nederlandse samenvatting within one day even when you have your own deadline soon. Thank you for helping me along the way especially in settling down a life in the Netherlands. Olga, thank you for our inspirational conversations from which I benefit strength and perseverance. Elly, thank you for offering the deal of proofreading all my manuscripts in exchange for lessons of Chinese language, which have been my enjoyable escaping moments from the frustrations in research. Min, Qian and Yifei, for our meaningful discussions of research and support for each other. Viktorija, for lightening up our office. Some of the many others I want to thank are prof. dr. Gang Cui, dr. Jessie Nixon, prof. dr. Jenny Doetjes, Jos Pacilly, dr. Kalinka Timmer, Gea Hakker, dr. Lesya Ganushchak, dr. Leticia Pablos, Mulugeta Tsegaye, prof. dr. Rint Sybesma, Yan Gu, dr. Xiaoqing Li, Yang Yang, prof. dr. Yufang Yang. I also want to thank everyone who dropped by my presentations at conferences and the anonymous reviewers of my submitted papers for the inspirational comments and suggestions.

Thank you, Kassandra, for being my truthful companion along this bumpy journey.

Last but not least, I want to thank my family: my parents - Huiling and Wentian, and my sister - Ying, for allowing me to pursue my own journey and supporting me unconditionally.

This dissertation was supported by grants from "Talent and Training China-Netherlands" program.

Table of Contents

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 Turennout, 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).

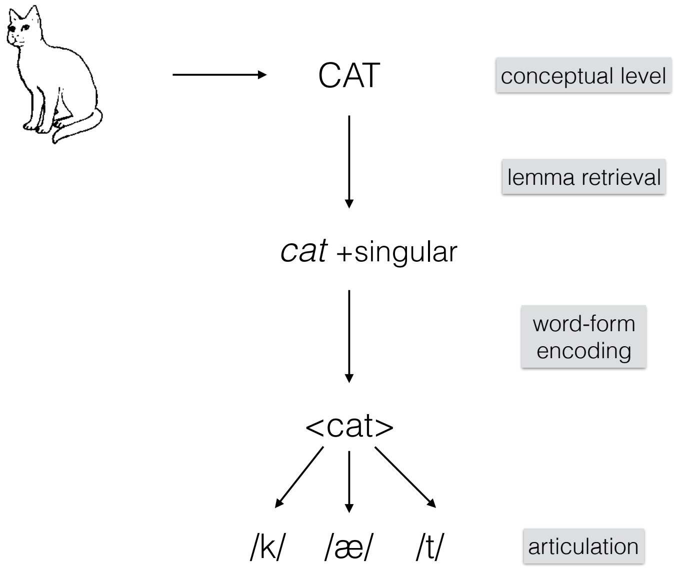


Figure 1.1 Stages of lexical access in WEAVER++ (adapted from Roelofs, 2000).

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 & Düngelhoff, 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).

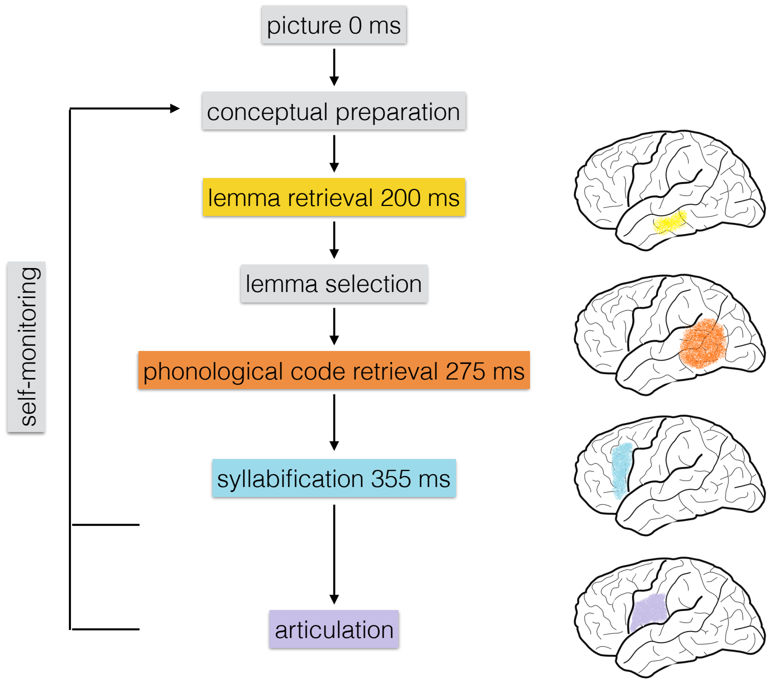


Figure 1.2 Schematic representation of the activation time course of brain areas involved in word production (adapted from Indefrey, 2011).

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 involoved in speech production in Mandarin Chinese.

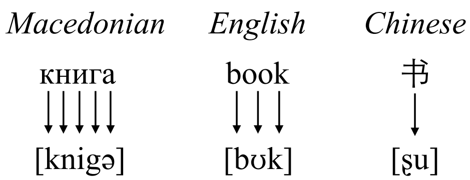


Figure 1.3 An illustration of the difference in depth of orthography in three exemplar languages.

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 comphrehesion, 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 & Düngelhoff, 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 Prato, 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.

Chapter 2

The contribution of orthography to spoken word production in blocked cyclic naming[1]

[1] A version of this chapter has been submitted for publication as Man Wang, Zeshu Shao, Antje S. Meyer, Yiya Chen, & Niels O. Schiller (submitted). The contribution of orthography to spoken word production in blocked cyclic naming.

Abstract

Does orthography contribute to spoken word production? Previous studies suggest that orthography is only involved in spoken word production when the orthographic representation is highly relevant, for instance, in reading aloud tasks. Using an adapted blocked cyclic naming paradigm, participants were asked to overtly name pictures that were presented repeatedly in semantically homogeneous, phonologically homogeneous, or heterogeneous blocks. On each trial, a written Chinese character that was either orthographically related or unrelated to the target was presented before the target picture. Chinese was selected as the target language because it is a language with relatively opaque mappings between orthography and phonology. We measured participants' speech onset latencies. Consistent with previous research, an inhibitory semantic blocking effect and a facilitative phonological blocking effect were found. More importantly, there was also an orthographic priming effect that was independent of both the semantic and the phonological effects. These findings suggest orthography contributes to speaking in a picture naming task, lending further support to the presence of orthographic priming in spoken word production, even in a language with a logographic script like Chinese.

2.1 Introduction

Language production, as an essential cognitive function in daily life, has drawn the attention of researchers for many years. Several influential models have been proposed to capture the underlying mechanisms of language production, in particular word production (e.g., Caramazza, 1997; Levelt, 1992, 1993; the spreading-activation model, Dell, 1990; Dell & O'Seaghdha, 1991, 1992; the WEAVER++ model, Levelt, Roelofs, & Meyer, 1999a, b; Roelofs, 1997; Roelofs & Meyer, 1998). Most of these models agree on the main stages involved in word production: (a) conceptualization of the intended message, (b) retrieval of the semantic and grammatical representations of the to-be-produced words (lemma), (c) word-form encoding and (d) articulation.

Most models postulate a modality-neutral lemma representation that is linked to phonological and orthographic representations of words (e.g. the WEAVER++ model). However, the Independent Network (IN) model (Caramazza, 1997; Rapp & Caramazza, 2002) assumes a modality-specific lexical representation, i.e. the phonological and orthographic representations of lexical items are independently connected to the semantic representation and they do not link to each other at the lexical level.

The modality-specific account, however, is challenged by evidence concerning the contribution of orthography to spoken word production. This issue has mostly been investigated using the form-preparation paradigm (Meyer, 1990), where participants first learn and memorize prompt-response word pairs (e.g. *sugar* - *COFFEE*). They are then presented with the probes and are asked to produce the corresponding response word. A facilitative effect has been reported when response words are phonologically related (e.g. *coffee, camel, cushion*) as compared to when they are unrelated (e.g. *coffee, scissors, giant*). Damian and Bowers (2003) reported that this facilitative effect in English is modulated by the consistency between phonology and orthography: The effect disappeared when phonology and orthography are not consistent (e.g. *camel* and *kennel*). This suggests the mandatory activation of orthography in speaking. However, this seems not to be the case in Dutch (Meyer, 1990, 1991; see Schiller, 2007 using a different paradigm), French (Alario, Perre, Castel, & Ziegler, 2007) or Chinese (Chen, Chen, & Dell, 2002).

Moreover, evidence suggests that activation of orthography in speaking is task dependent. For instance, orthographic inconsistency showed an inhibitory effect in a reading task but not in picture naming, word generation or associative naming, such as *contract, kanon, konijn (contract, cannon, rabbit)*, compared to *contract, colbert, cadeau (contract, jacket, present)* in Dutch (Roelofs, 2006; see Bi, Wei, Janssen, & Han, 2009 for similar findings in Chinese). Using the picture-word interference paradigm, where a written distractor word is displayed simultaneously with a picture, orthographically-related distractors facilitate picture naming in Mandarin Chinese (e.g., Zhang, Chen, Weekes, & Yang, 2009; Zhang & Weekes, 2009; Zhao, La Heij, & Schiller, 2012). Taken together, these results suggest that orthography only influences speech production when it is highly relevant to the task.

The controversial evidence for the involvement of orthography in speaking is possibly also affected by the degree of transparency of orthography-to-phonology mappings (Roelofs, 2006). In languages with an alphabetic script, orthography corresponds directly to phonology (so when sound overlaps, orthography tends to also overlap), and therefore effects of phonology and orthography are often confounded. Chinese, as a language with relatively opaque mapping between phonology and orthography, can serve as an appropriate target language to dissociate phonological and orthographic effects, because it is easy to find items with only phonological overlap or with only orthographic overlap.

Notably, given different phonology-orthography mapping rules across languages, the level of interactions between phonology and orthography may also vary between languages with alphabetic and non-alphabetic scripts. Qu and colleagues (Qu, Damian, & Li, 2016) proposed that, although in both languages the semantic system activates phonology and orthography respectively and a modality-specific lexicon is activated accordingly, the link between phonology and orthography at the sublexical level is distinct (see Figure 2.1).

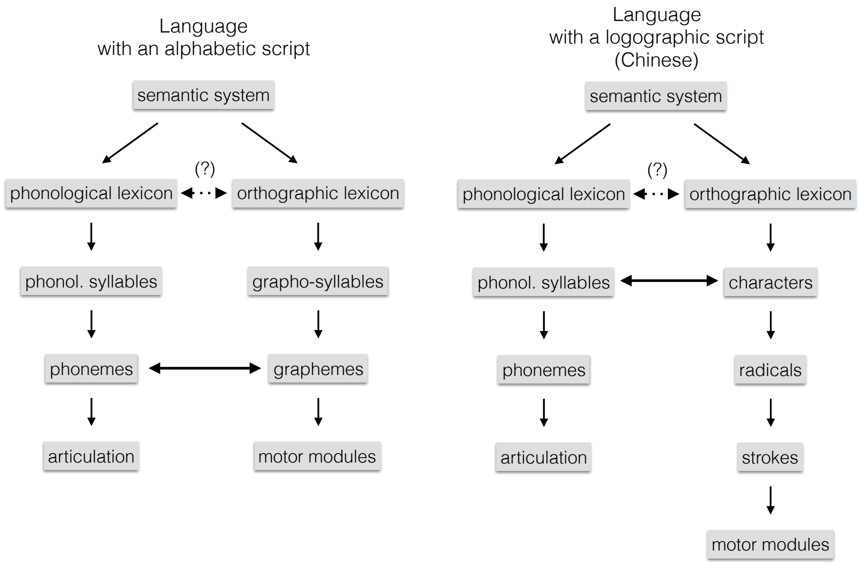


Figure 2.1 A model of word production system for speaking and writing in languages with alphabetic and non-alphabetic scripts (adapted from Qu et al., 2016).

To examine the effect of orthography in different stages during speaking, we used an adapted blocked cyclic naming paradigm. Blocked cyclic naming has mainly been used to study language production. In this paradigm, participants are required to name a series of pictures repeatedly in cycles where either targets belong to the same semantic category (hereafter semantically homogeneous block), like *eye, nose, arm, shoulder*, or targets overlap in their phonological onset segments (hereafter phonologically homogeneous block), like *bean, bell, boot, bowl.* A control condition is provided by grouping together unrelated items (hereafter heterogeneous block), like *eye, desk, goat, sweater.* Participants are often slower in naming pictures in the semantically homogeneous blocks than in the heterogeneous blocks, i.e. the semantic blocking effect[2] (e.g., Belke, 2013; Belke, Meyer, & Damian, 2005; Belke & Stielow, 2013; Damian, Vigliocco, & Levelt, 2001; but see Navarrete, Del Prato, & Mahon, 2012; Navarrete, Del Prato, Peressotti, & Mahon, 2014) and faster in the phonologically homogeneous blocks than in the heterogeneous blocks, i.e. the phonological facilitation effect (e.g., Damian, 2003; Damian & Stadthagen-Gonzalez, 2009; but see Damian & Dumay, 2009).

In the present study, we combined priming with the blocked cyclic naming paradigm. The primes were written Chinese characters that were either orthographically related or unrelated to the target picture name. Target pictures were repeated in the semantically homogeneous, phonologically homogeneous, or heterogeneous blocks[3]. This design allows us to examine the interplay between orthographic encoding and semantic retrieval or phonological encoding in speaking.

According to the literature, we expect to observe the semantic interference effect when comparing semantically homogeneous and heterogeneous blocks. Notably, in phonologically homogeneous blocks, the phonological forms of target pictures shared the first two segments (smaller than a syllable) in terms

---

[2] In the present study, we will refer to this slowing-down effect observed in the blocked cyclic naming paradigm as the *semantic blocking effect* to differentiate it from semantic interference effects in the cumulative semantic interference paradigm (Costa, Strijkers, Martin, & Thierry, 2009; Howard et al., 2006; Navarrete, Mahon, & Caramazza, 2010) or the picture-word interference paradigm (e.g., Glaser & Düngelhoff, 1984; Schriefers, Meyer, & Levelt, 1990).

[3] It would be ideal to have an orthographically homogeneous block as well. However, it is not feasible to find picture names with orthographic (but not phonological or semantic) overlap while satisfying all other criteria of the picture selection (e.g., high frequency picture names, familiar objects, high naming agreement, low visual complexity, etc.).

of pinyin, i.e. the phonetic notation of Chinese characters (e.g., */bi/, /bian/, /biao/*). It has been debated over whether the unit of phonological encoding in Chinese is the syllable (Chen et al., 2002; O'Seaghdha, Chen, & Chen, 2010) or smaller phonological elements (e.g., Qu, Damian, & Kazanina, 2012; Verdonschot et al., 2015). If the phonological effect is observed in the present study, this would provide new evidence for the size of the phonological encoding unit in word production in Chinese. Most importantly, we expect that orthographic overlap should facilitate spoken word production even when orthographic information is not highly relevant for spoken word production (Qu et al., 2016). Moreover, we are interested in whether and how the priming effect of orthography interacts with semantic and/or phonological effects in speaking. If orthography influences spoken word production via the lexical-semantic pathway (i.e. a link from the orthographic lexicon to semantic representation; as proposed in Zhang & Weekes, 2009), we should observe an attenuated semantic interference effect in the orthographically-related condition in the semantically homogeneous blocks compared to the heterogeneous blocks, because orthographic primes spread activation to the semantic representations of targets. If orthography influences phonological encoding, we should find an interaction between the effects of orthography and phonology.

2.2 Methods

2.2.1 Participants. Twenty-five native speakers of Chinese (10 males, mean age = 26.5 years, SD = 4.0 years) studying in the Netherlands gave informed consent and participated in the experiment. All participants had normal or corrected-to-normal vision and no history of language deficits. They received 7 euros for their participation.

2.2.2 Materials and design. Thirty-two line drawings of common objects were selected from the CRL-IPNP (CRL International Picture Naming Project, Bates et al., 2000) and the standardized Snodgrass and Vanderwart picture database (Snodgrass & Vanderwart, 1980). Twenty-nine picture names were disyllabic and three were tri-syllabic. Pictures were standardized to 300 × 300 pixels and appeared in the center of the screen as black drawings on a white background.

Half of the pictures were combined to create four semantically homogeneous blocks (see Appendix I), with four pictures belonging to the same semantic category in each block.

The other half of the pictures were combined to create four phonologically homogeneous blocks, with four pictures in each block (see Appendix I). In each block, target names of pictures shared the first two phonological segments.

To form the heterogeneous blocks, sixteen pictures were randomly selected from the phonologically or semantically homogeneous blocks and grouped into four blocks, where picture names were neither phonologically nor semantically related.

Table 2.1 The pictures used in the semantically homogeneous, phonologically homogeneous and heterogeneous blocks are comparable in terms of naming agreement, word frequency, age of acquisition (AOA) and visual complexity (see Bates et al., 2000 for the details of the norms).

	$F(2, 32)$	p-value
naming agreement	1.20	0.31
word frequency	2.07	0.14
AOA	0.64	0.54
visual complexity	0.27	0.77

Each picture was paired with one orthographically-related (e.g. 层, ceng2, 'layer'), or unrelated character (e.g., 自, zi4, 'self') as primes for the same target (e.g. 犀, xi1, 'rhinoceros'). Orthographic condition was also blocked.

The experiment had a within-participants design. In each block, all prime-picture pairs were repeated four times in a cyclic manner. In total, each participant named 384 pictures. The prime-picture pairs in each cycle were presented in a pseudo-randomized manner such that the same picture did not appear in the same order in two consecutive cycles or three consecutive trials and the same block condition nor prime condition appeared in two consecutive blocks. The stimulus lists were counterbalanced across participants.

2.2.3 Procedure and apparatus. Participants were seated approximately 50 cm away from a computer screen in a soundproof booth. Stimuli were presented using E-prime 2.0 and the reaction times (RTs; i.e. the speech onset latencies) were measured online by a voice-key connected with a PST serial response box. The participants' vocal responses were also recorded. Mistriggered RTs were corrected manually in Praat based on the recordings. Speech errors were first manually coded during the experiment and then double-checked against the recordings.

Before the experiment, participants were familiarized with the pictures used in the experiment. Each picture was presented once in the center of the computer screen for 2 s in a randomized order, and participants were asked to name the pictures. Participants were corrected if they used a non-dominant name. On each practice trial, a fixation cross appeared in the center of the screen for 500 ms, followed by an 'X' for 75 ms. Then the target picture appeared and lasted until the voice-key was triggered or a 2 s limit was exceeded, followed by a blank screen for 1 s.

The procedure on the experimental trials was the same as for the practice trials except that the "X" was replaced by a Chinese character prime. There was a warm-up session preceding each experimental list, consisting of four prime-picture pairs, which were not included in the experimental stimuli. There were self-paced pauses between blocks. The whole session lasted about 20 minutes.

2.2.4 Data analysis. Incorrect and disfluent responses were considered as errors and excluded from the RT analysis. The error rate (3.07%) was considered too low to warrant analysis. RTs beyond three SDs from the mean (by participant) were considered as outliers (1.97%) and were excluded. The naming RTs showed a skewed distribution and therefore were log-transformed (base 10). The naming RTs showed a normal distribution after log-transformation. The log-transformed RTs (9,116 data points) were analyzed using mixed-effects modeling in R (version 3.1.0; R Core Team, 2014) using the 'lmerTest' package (Kuznetsova, Brockhoff, & Christensen, 2015). Following a maximal-model approach (Barr, Levy, Scheepers, & Tily, 2013), the initial model was built with two fixed factors: block condition (three levels: semantically homogeneous, phonologically homogeneous, and heterogeneous) and prime condition (two levels: orthographically related and unrelated), two random intercepts: participants and target pictures and one control variable: presentation cycle (cycle 1, 2, 3 and 4). Interactions between fixed factors, by-participant random slope of the fixed factors, and the by-item random slope of the control variable were also tested.

2.3 Results

Figure 2.2 and Table 2.2 summarize the results. First, compared to the heterogeneous blocks (mean = 590 ms, SD = 67 ms), RTs were significantly longer in the semantically homogeneous blocks (mean = 619 ms, SD = 66 ms), and significantly shorter in the phonologically homogenous blocks (mean = 569 ms, SD = 63 ms). There was also a significant effect of presentation cycle.

Critically, we obtained a main effect of orthographic prime condition such that the RTs were shorter by 12 ms in the orthographically-related condition (mean = 586 ms, SD = 64 ms) than in the unrelated condition (mean = 598 ms, SD = 65 ms). This suggests that the orthographically-related primes facilitated picture naming (see Figure 2.2). No interactions between prime type and block condition were found.

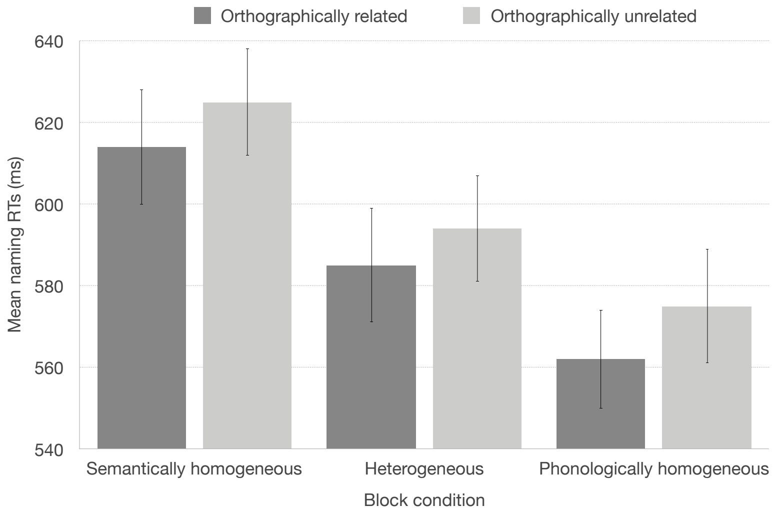


Figure 2.2. Mean naming RTs (in ms) in semantically homogeneous, heterogeneous, and phonologically homogeneous conditions, for orthographically-related (dark gray) and unrelated (light gray) conditions. The error bars represent the positive and negative standard errors of the mean in each condition.

Table 2.2 Results summary: coefficient estimates, standard errors (SE) and *t*-values in the final model.

	Coefficient Estimate	SE	*t*-value	*p*-value
Intercept	6.460658	0.026097	247.57	< 0.0001
Orthographically-related primes	-0.017603	0.005769	-3.05	0.0026
Phonological block	-0.018505	0.007130	-2.60	0.0095
Semantic block	0.031882	0.006685	4.77	< 0.0001
Cycle	-0.036111	0.004145	-8.71	< 0.0001
Orthographically : Phonologically related	-0.002964	0.008060	-0.37	0.7131
Orthographically : Semantically related	-0.001378	0.008145	-0.17	0.8658

2.4 Discussion

To the best of our knowledge, the present study is the first to show the semantic interference effect in the blocked cyclic naming paradigm, in Chinese. The magnitude of the semantic effect in the present study (34 ms) is similar to previous research (e.g., 30 ms for English, in Howard, Nickels, Coltheart, & Cole-Virtue, 2006, and 27 ms for Dutch in Shao, Roelofs, Martin, & Meyer, 2015). The consistency of the semantic interference effect across languages suggests that the semantic effect is unlikely to be driven by a language-specific mechanism. The semantic interference effect is interpreted as reflecting competition during lexical selection when multiple candidates from the same semantic category are activated (e.g., Belke, 2013; Belke et al., 2005; Belke & Stielow, 2013; Damian et al., 2001; but see Navarrete et al., 2014 for an alternative interpretation).

Secondly, we obtained a phonological facilitation effect: Participants responded faster in the phonologically homogeneous blocks than in the heterogeneous blocks. This result is in line with the findings reported in previous studies (e.g., Damian, 2003; Damian & Stadthagen-Gonzalez, 2009; Roelofs, 1999). The phonological facilitation effect was obtained with sub-syllabic overlap, suggesting a phonological encoding unit smaller than the syllable in Chinese (Qu et al., 2012; Verdonschot et al., 2015). The magnitude of the phonological facilitation effect (22 ms) on naming is comparable to that obtained in previous studies (17 ms and 27 ms in Qu et al., 2016 in Chinese and on average 31 ms in Damian and Martin, 1999 in English). Note that most participants recruited in the present study speak either Dutch or English as their second language and had lived in the Netherlands for at least two months. It is therefore possible that their exposure to languages with an alphabetic script had influenced the size of phonological encoding units in Chinese (see Verdonschot, Nakayama, Zhang, Tamaoka, & Schiller, 2013).

Most importantly, we found an orthographic facilitation effect, indicating that the activation of an orthographic representation can facilitate lexical access in spoken word production. The effect was present from the first cycle in the blocked cyclic naming paradigm with orthographic priming, and thus could not have originated from a learning phase (see Alario et al., 2007). This result is consistent with previous studies that found that orthography could influence the production of spoken words (e.g. Damian & Bowers, 2005). The orthographic facilitation effect suggests orthographic relatedness can contribute to speaking in Mandarin Chinese.

It is important to note that the size of orthographic facilitation effect was similar in the semantically and phonologically homogeneous blocks and it did not interact with the semantic or phonological effects. This suggests that orthography is unlikely to influence lexical access via the lexical-semantic pathway (see Zhang & Weekes, 2009) or via spreading activation from the activated phonological form. Rather, orthography has an independent impact on spoken word production. Where, then, does the orthographic priming effect arise? One possibility is that orthography affects spoken word production via an orthography-to-phonology link at the sublexical level, compatible to the Qu et al. (2016) model.

In summary, using Chinese, a language with relatively opaque mappings between orthography and phonology, we found clear evidence for the contribution of orthography to spoken word production even when orthographic information is not highly relevant for production. In addition, we have found the semantic blocking effect in Chinese and contributed to the understanding of the phonological unit in Chinese spoken word production. Future studies and models of spoken word production should take these results into account.

Acknowledgments

This research was supported by grants from "Talent and Training China-Netherlands" program. We thank Jos Pacilly for the modified Praat script for checking speaking reaction times. The article benefited from discussions with Markus Damian (Bristol) and feedback obtained during a poster presentation at the International Workshop on Language Production in Geneva (Switzerland), July 2014. We thank Elly Dutton for proofreading this manuscript.

Chapter 3

The time course of speech production revisited: No early orthographic effect, even in Mandarin Chinese[4]

[4] A version of this chapter has been submitted for publication as Man Wang, Yiya Chen, Minghu Jiang, & Niels O. Schiller (submitted). The time course of speech production revisted: No early orthographic effect, even in Mandarin Chinese.

Abstract

Most psycholinguistic models of speech production agree on an earlier semantic processing stage and a later word-form encoding stage. Using a language with a logographic script, Mandarin Chinese, Zhang and Weekes (2009) reported an early effect of orthography in a picture-word-interference study and suggested that orthography can affect speech production via a lexical-semantic pathway at an early stage. This early orthographic effect without co-occurrence of phonological effect, however, was not replicated (Zhao, La Heij, & Schiller, 2012). The present study aimed to shed light on the contradictory results and further tap into the potential interaction and time course of orthography and semantic processing. Experiment 1 re-investigated the orthographic effect on picture naming. The results demonstrated a semantic interference effect at negative SOAs while orthographic relatedness facilitated picture naming at a positive SOA. No interaction between semantic and orthographic relatedness was found. The results thus replicated Zhao et al. (2012) with a late effect of orthography. Given that in both Experiment 1 and previous studies, complex Chinese characters were used as stimuli with sub-parts indicating either the sound or the meaning of the whole characters, the different results with respect to Zhang and Weekes (2009) could have resulted from varying degrees of overlap between orthographic and either phonological or semantic information. Experiment 2 therefore used simplex Chinese characters so as to clearly dissociate the semantic and phonological representations from orthography. The results revealed an orthographic effect but only at a similar point in time as the phonological effect, both of which followed the semantic effect. Taken together, our results raise doubts about the role of orthography at the conceptual level of speech planning and lend further support to a two-step model of speech production.

3.1 Introduction

An important issue in psycholinguistic research is the extent to which psycholinguistic models are capable of accounting for cross-linguistic differences. Models of speech production 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). Previous studies have reported that orthographic relatedness modulates the speech production response latencies (Lupker, 1982; Posnansky & Rayner 1978; Underwood & Briggs, 1984). However, models of speech production have been mainly based on evidence from West Germanic languages, where orthographic and phonological forms are less clearly distinguished. For instance, the WEAVER++ model postulates a modality-neutral lemma representation where orthography is not specified (Levelt, Roelofs, & Meyer, 1999a, b; Roelofs, 1992; Roelofs & Meyer, 1998). Alternatively, the Independent Network model (Caramazza, 1997; Rapp & Caramazza, 2002) postulates a modality-specific representation in language production with the semantic representation activating the phonological representation of the lexicon in speech production and orthographic representation in written word production. In other words, the Independent Network model recognizes the role of the orthographic representation but posits that it only affects written word production.

It is difficult to tease apart orthography and phonology in languages with an alphabetic script because the correspondence between grapheme and phoneme is relatively transparent with some showing very consistent mapping (as in Serbo-Croatian) but others relatively less consistent mapping (as in English) (Katz & Frost, 1992). By contrast, languages with a logographic script show a highly arbitrary grapheme-to-phoneme correspondence. Take Mandarin

Chinese as an example; the basic unit of the writing system is a logographic character, and one character usually corresponds to a syllable. The number of possible syllables in Mandarin Chinese is limited, i.e. about 400 syllables excluding lexical tones or about 1,300 syllables including tones (Duanmu, 2002). As a consequence, there is a large number of homophones, with the result that orthography plays a crucial role in distinguishing homophones. It is therefore possible that in languages with a logographic script such as Mandarin Chinese, orthography plays a different role in speech production compared to languages with an alphabetic script.

Attempts to address the separate roles of orthography and phonology in speech production have been made in English (Damian & Bowers, 2009; Lupker, 1982; Posnansky & Rayner, 1978) using the picture-word interference paradigm (e.g., Lupker, 1979; Rosinski, Golinkoff, & Kukish, 1975). In this paradigm, participants are asked to name pictures while ignoring superimposed distractor words. It is found that distractor words that belong to the same semantic category as the target interfere with picture naming and phonologically-related distractors facilitate picture naming (e.g., Starreveld, 2000; Starreveld & La Heij, 1995, 1996; see Glaser, 1992; MacLeod, 1991 for a review of the paradigm). When the distractors are both orthographically and phonologically related to the picture name, the facilitation effect is stronger compared to pure phonological relatedness (e.g., Lupker, 1982; Posnansky & Rayner 1978; Underwood & Briggs, 1984). For instance, naming the picture of a *chair* was faster with the distractor *air* (55 ms) or *bear* (23 ms), compared to an unrelated condition, from which the facilitation effect size was derived (32 ms) and attributed to orthographic overlap (Lupker, 1982). However, Damian and Bowers (2009) found that 'extra' orthography alone did not modulate the facilitation effect when distractors were presented in the auditory format instead of the visual modality. Therefore, the presence of a pure orthographic effect in speech production has remained unclear.

Two factors may have contributed to the discrepancy in the results of the studies based on English stimuli. One factor is the limited number of word pairs that can dissociate orthography and phonology in English (e.g. *bear – year*). The other factor is that the role of orthography was often not examined independently but rather tested by a subtraction approach (the effect of phonological and orthographic relatedness minus the effect of phonological relatedness; e.g., Lupker, 1982; Posnansky & Rayner 1978; Underwood & Briggs, 1984). Damian and Bowers (2009) pointed out that one of the limitations of using English words as stimuli is that the distractors in the orthographically unrelated condition were only orthographically "less similar". Consequently, this might have "underestimated the potential contribution of spelling" (Damian & Bowers, 2009, p. 595).

Mandarin Chinese provides an ideal testing ground to tease apart the role of orthography and phonology in speech production. As we mentioned earlier, it has a logographic writing system that can easily dissociate phonology and orthography. Each syllable in Mandarin Chinese contains segmental information and a lexical tone, and is represented by a single character that comprises one or more sub-elements, known as 'radicals'. A semantic radical is a sub-element of a Chinese character that conveys semantic information about the character, while a phonetic radical conveys phonological information about the character. For example, 锤 (chui2, 'hammer') (here *chui* is the 'pinyin' transcription of the Mandarin syllable, and *2* indicates Lexical Tone 2) is a complex character where the left part is a semantic radical 钅 indicating that it is related to metal, and the right part is the phonetic radical 垂 (chui2, 'suspend') indicating the sound of the character 锤 (chui2, 'hammer'). Some characters, however, contain only one element (henceforth 'simplex' characters). For example, 羊 (yang2, 'sheep') is a simplex character which cannot be decomposed into sub-parts. It can be seen, then, that Chinese characters may overlap in phonology but not in orthography, and vice versa. For example, simplex 羊 (yang2, 'sheep') and 央 (yang1, 'center') are only phonologically related (i.e. overlapping at the segmental level *yang* although differing in lexical tones), while 羊 (yang2, 'sheep') and 半 (ban4, 'half') are orthographically related but have no phonological overlap (i.e. neither in segment nor in tone).

Independent orthographic and phonological facilitation effects have been reported in studies using Mandarin Chinese stimuli (Bi, Xu, & Caramazza, 2009; Zhang, Chen, Weekes, & Yang, 2009; Zhang & Weekes, 2009; Zhao, La Heij, & Schiller, 2012). Nevertheless, studies that have manipulated the stimulus onset asynchrony (SOA) have yielded mixed results regarding the temporal locus of the orthographic effect (Zhang et al., 2009; Zhang & Weekes, 2009; Zhao et al., 2012). Using the picture-word interference paradigm, Zhang and colleagues (Zhang et al., 2009; Zhang & Weekes, 2009) reported orthographic effects with the negative SOAs (-150 ms and -100 ms) without co-occurrence of any phonological effect, which led them to claim that sharing orthography might activate the target concept via the lexical-semantic pathway (Link A in Figure 3.1) and facilitate the target name retrieval at an earlier stage compared to the phonological effect. However, the results were not replicated by Zhao et al. (2012). Instead, their results demonstrated that orthographically and phonologically related distractors both facilitated picture naming at a similar stage, i.e. the word-form encoding stage of speech production.

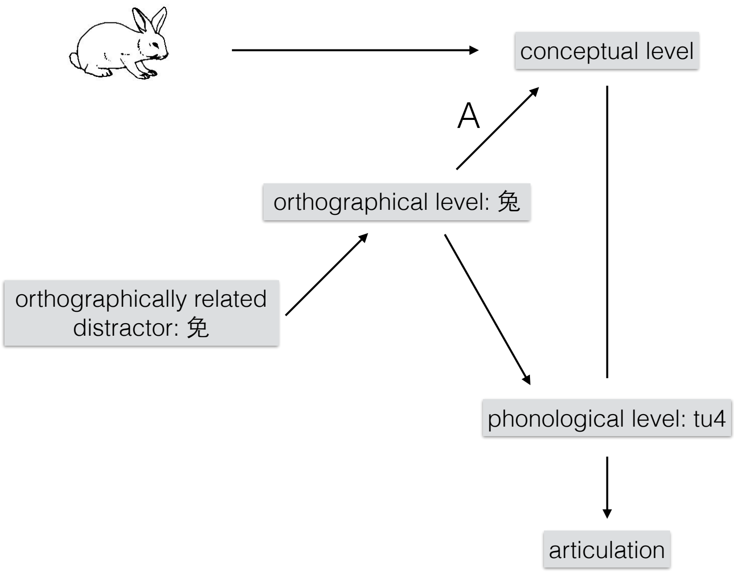


Figure 3.1 The model of overt picture naming with distractors in Chinese (adapted from Zhang & Weekes, 2009; Zhao et al., 2012).

In addition to the lack of consensus in the literature regarding the time course of the orthographic effect on picture naming, another issue that has not been explicitly addressed in the existing literature, is whether orthographically-related distractors affect speech production by interacting with the related semantic representation of the target word. The goal of Experiment 1 of the present study was therefore two-fold. First, we were interested in resolving the controversy whether orthographically-related distractors affect speech production via a lexical-semantic pathway independent of the phonological effect. Second, we were interested in whether orthographically-related distractors affect speech production by interacting with semantics. To this end, we employ a full factorial design including all four possible conditions of semantic and orthographic overlap: semantically and orthographically related, semantically related but orthographically unrelated, orthographically related but semantically unrelated, and unrelated. We use the picture-word interference paradigm with SOAs ranging from negative to positive values to cover the process before and after the activation of the target lemma respectively (see Schriefers et al., 1990; Zhang & Weekes, 2009; Zhao et al., 2012). A more refined increment (75 ms) is employed (instead of 100 ms as in Zhang & Weekes, 2009) to increase the sensitivity of detecting the hypothesized effects. If orthography facilitates speech production at the conceptual level, as claimed in Zhang and Weekes (2009), we expect an orthographic effect at negative SOAs, possibly with the same temporal locus as that of the semantic effect (Zhang & Weekes, 2009) or interacts with the semantic effect.

As we noted earlier, in Mandarin Chinese, simplex characters and complex characters have distinctive structural properties. Given that we used complex characters in Experiment 1 to test possible interactions between semantics and orthography, we also designed Experiment 2 with only simplex-character stimuli to further disentangle orthography from semantics and phonology. Such a design allows us to zoom into the orthographic effect as well as semantics and phonological effects on speech production without having to worry about the possible overlap between orthography and semantics or phonology. The time course of these effects can then be more clearly teased apart.

3.2 Experiment 1

3.2.1 Methods

3.2.1.1 Participants. Twenty native Mandarin speakers (5 male; average age = 27.4 years; SD = 2.41 years) studying in the Netherlands were paid for their participation. All participants signed a letter of informed consent, had normal or corrected-to-normal vision and none had any language impairments.

3.2.1.2 Materials and design. Twenty black-and-white line drawings from the International Picture Naming Project (Bates et al., 2003) and Snodgrass and Vanderwart (1980) databases, or drawn similarly, corresponding to complex character names in Mandarin Chinese (either monosyllabic N = 7 or disyllabic N = 13) were selected as target pictures. Each picture was presented with four types of monosyllabic distractors: a) semantically and orthographically related (S+O+); b) semantically related but orthographically unrelated (S+O-); c) orthographically related but semantically unrelated (S-O+); d) semantically and orthographically unrelated (S-O-). Ten other pictures corresponding to monosyllabic or disyllabic names were selected from the same databases to serve as fillers.

All the distractors were phonologically unrelated to the targets. The distractors in the four conditions were comparable in terms of word frequency, $F(3, 76) < 1$ (calculated with the log frequency of words in the SUBTLEX-CH database; Cai & Brysbaert, 2010) and visual complexity (number of strokes), $F(3, 76) = 1.655$, $p > .05$. Orthographic relatedness was operationalized by overlapping in one radical of the characters (e.g. 猫, mao1, 'cat' and 狗, gou3, 'dog' which overlap in the radical 犭). Fourteen native Mandarin speakers were asked to rate the semantic relatedness of word pairs with one distractor word and its corresponding target word on a 1-7 scale, with the higher score indicating stronger relatedness. The average rating scores per participant were then submitted to Wilcoxon Signed-Rank tests. The rating scores differed significantly between semantically related and unrelated word pairs, $Z = -3.9$, $p < .0001$. The semantic relatedness did not differ between S+O+ and S+O-, $Z = -1.9$, $p > .05$ or S-O+ and S-O-, $Z = -1.4$, $p > .05$.

The design included two factors: Distractor Type (S+O+, S+O-, S-O+, S-O-) and SOA (-150 ms, -75 ms, 0 ms and 75 ms). Each participant received 30 pictures × 4 Distractor Types × 4 SOAs = 480 trials in total in a pseudo-random order such that the same picture did not re-occur within three consecutive trials. The trials were blocked by SOA. The sequence of the blocks was counterbalanced across participants.

3.2.1.3 Apparatus and procedure. Before the experiment there was a familiarization and practice session. The participants were first shown all the pictures with their names underneath, and were then asked to name the pictures without their names presented. Incorrect answers were corrected.

Each trial in the experimental sessions consisted of: a fixation (300 ms); a blank screen (200 ms); the first stimulus that was either the target picture (350 by 350 pixels) or the distractor depending on the SOA (Arial Unicode MS, 48 point size); followed by the second stimulus (again either target picture or distractor). The stimuli lasted until the voice-key was triggered or a 2 s limit was exceeded, followed by another blank screen (500 ms). There was a self-paced pause between every two blocks.

The stimuli were presented using the software E-prime 2.0 and reaction times were recorded online by a voice-key connected with a PST serial response box. Incorrectly triggered voice-key responses were corrected manually using the program CheckVocal (Protopapas, 2007). Errors were firstly manually coded on-line and then double-checked based on the voice recordings.

3.2.2 Results and discussion

Errors (3.41% of all 6,400 data points; including incorrect and disfluent responses) and outliers (1.17%; shorter than 300 ms and longer than 1,300 ms) were excluded from further analysis. Error rates were very low and thus considered not informative enough for further statistical analysis. The naming latencies showed a skewed distribution and were therefore log-transformed (base 10). The log-transformed naming latencies (6,107 data points) were submitted to the mixed-effects modelling in R (version 3.1.0; R Core Team, 2014) as the dependent variable.

Table 3.1 The average naming latencies (ms) and percentage errors (in parentheses) for each condition in Experiment 1.

	Distractor type			
	Semantically related		Semantically unrelated	
SOA (ms)	Orthographically related	Orthographically unrelated	Orthographically related	Orthographically unrelated
-150	708 (.20)	713 (.22)	698 (.17)	692 (.19)
-75	719 (.22)	738 (.20)	712 (.19)	713 (.17)
0	744 (.13)	749 (.22)	724 (.27)	728 (.30)
75	730 (.25)	750 (.34)	725 (.16)	733 (.19)

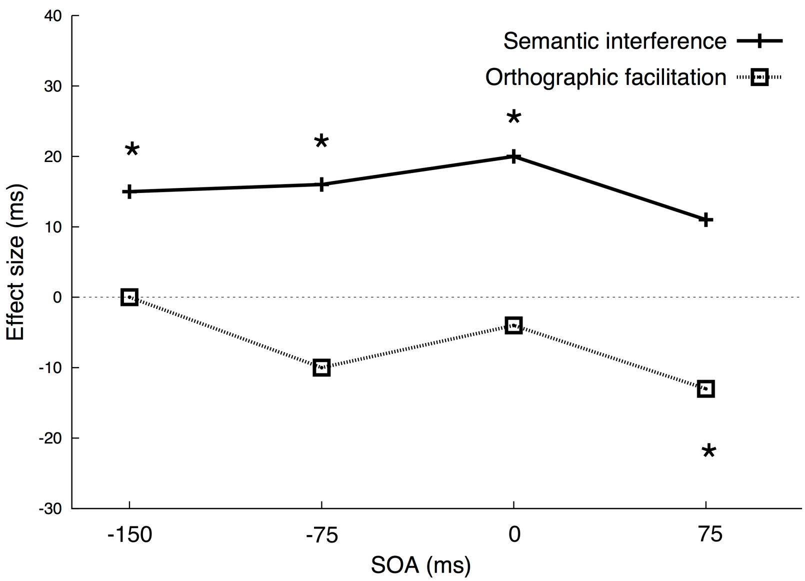


Figure 3.2 The main effects of semantic and orthographic distractors on picture naming in Experiment 1.

The initial statistical model was built using the 'lmer4' package (Bates, Maechler, Bolker, & Walker, 2014) following a maximal-model approach (Barr, Levy, Scheeper, & Tily, 2013). The initial model included three fixed predictors: semantic relatedness, orthographic relatedness and SOA, two-way interactions between distractor type (semantic and orthographic relatedness) and SOA, two random intercepts: participant and target picture, and the random slopes of fixed predictors by participant. The model failed to converge so the least variable random slope (the random slope of orthographic relatedness by participant) was removed. The interaction between orthographic relatedness and SOA was significant, $t > 1.65$ (one-tail; based on Zhang et al., 2009; Zhang & Weekes, 2009; Zhao et al., 2012). The data were then divided into four subsets per SOA. Separate models were built with semantic relatedness and orthographic relatedness as the fixed predictors, the random intercepts: the participant and target picture, and the random slopes of fixed predictors by participant. The *p*-values were obtained using the 'pbkrtest' package (Halekoh & Højsgaard, 2014).

Table 3.2 The results summary: coefficient estimates, standard errors (SE), t-values and *p*-values for the effect of distractor type in each SOA condition in Experiment 1. (significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 "." 0.1 "" 1)

SOA (ms)	Distractor Type	Coefficient Estimate	SE	*t*-Value	*p*-Value
-150	Intercept	6.5274	0.0291	224.3	
	Semantic relatedness	0.0204	0.0079	2.6	0.014*
	Orthographic relatedness	0.0015	0.0078	0.2	> 0.05
-75	Intercept	6.5598	0.0238	275.7	
	Semantic relatedness	0.0206	0.0083	2.5	0.018*
	Orthographic relatedness	-0.0136	0.0086	-1.6	> 0.05
0	Intercept	6.5764	0.0278	236.4	
	Semantic relatedness	0.0265	0.0084	3.2	0.003**
	Orthographic relatedness	-0.0099	0.0093	-1.1	> 0.05
75	Intercept	6.5827	0.0256	256.9	
	Semantic relatedness	0.0161	0.0083	1.95	> 0.05
	Orthographic relatedness	-0.0188	0.0085	-2.2	0.035*

When SOA was -150 ms, -75 ms or 0 ms, there was a significant effect of semantic interference (+15 ms, +16 ms and +20 ms respectively). Naming latencies with semantically related distractors were significantly longer than those with semantically unrelated distractors (see, e.g., La Heij, 1988; Levelt et al., 1999a; 1999b; Roelofs, 2003; but see also, e.g., Finkbeiner & Caramazza, 2006; Finkbeiner, Gollan, & Caramazza, 2006; Mahon, Costa, Peterson, Vargas,

& Caramazza, 2007; Miozzo & Caramazza, 2003 for accounts of the semantic effect). There was a significant effect of orthographic facilitation when SOA was 75 ms (difference of -13 ms). The semantic effect did not reach significance at SOA of 75 ms. The interaction between the semantic and orthographic factors did not reach significance at any SOA.

The semantic interference effect was shown at negative SOAs. This result is compatible with previous research using the picture-word interference paradigm in both languages with an alphabetic script and languages with a logographic script (e.g., Lupker, 1982; Zhang & Weekes, 2009; Zhang et al., 2009).

Critically, we did not observe an early orthographic effect or an interaction between orthographic relatedness and semantic relatedness at negative SOAs. Instead, the orthographic effect was only demonstrated with the positive SOA (i.e. 75 ms), suggesting the orthographic relatedness only affects the picture naming process after lemma retrieval, possibly at the word-form processing stage. This result did not confirm the necessity to reconstruct the speech production model regarding the orthographic effect, as suggested by Zhang and Weekes (2009).

It is worth noting that the significant semantic and orthographic effects have distinctive temporal loci without any overlap at the specified SOAs. That is, the semantic interference effect was only found at negative SOAs and orthographic facilitation at positive SOAs. This pattern is similar to the pattern of results in Schriefers et al. (1990), suggesting a two-step model of speech production that distinguishes meaning and form processing (but see e.g. Dell, Schwartz, Martin, Saffran, & Gagnon, 1997 for an interactive two-step model).

Furthermore, the effect sizes of the semantic interference and orthographic facilitation were comparable to those in Zhang and Weekes (2009) but smaller than Zhao et al. (2012). In contrast to Zhang and Weekes (2009), there was only a numerical difference between the orthographically related and the unrelated conditions at negative SOAs (-10 ms at SOA -75 ms and -4 ms at SOA 0 ms). Moreover, the size of the orthographic facilitation effect obtained at SOA 75 ms was relatively small (-13 ms) with a p-value of .035. There is a possibility that the current design is not sensitive enough to obtain a robust orthographic effect. For instance, the orthographic relatedness represented by sharing one radical (e.g. 碗, wan3, 'bowl' and 矿, kuang4, 'mine' share the radical 石, shi2, 'stone') may not be salient enough to facilitate picture naming. It has been discussed in the Chinese character literature that the characters are likely to be processed as a whole, in line with a holistic processing view (e.g., evidence from Cheng, 1981; Tzeng, Hung, Cotton, & Wang, 1979; Yu, Feng, Cao, & Li, 1990; but see evidence from Feldman & Siok, 1999; Yeh & Li, 2004 for an analytic view). Consequently, it is possible that the partial overlap is not perceptually processed individually and therefore did not produce an orthographic effect at negative SOAs.

Experiment 2 was designed to tap into the time course of the orthographic effect using simplex characters with orthographic relatedness implemented as the whole-character orthographic similarity. Another advantage of using simplex characters is that we can avoid implicit confounding effects of orthography and phonology or semantic information.

3.3 Experiment 2

3.3.1 Methods

3.3.1.1 Participants. Sixty-eight native Mandarin speakers (30 male; average age = 21.6 years; SD = 2.19 years) living in Beijing, China were paid for their participation in the experiment. All participants signed a letter of informed consent, had normal or corrected-to-normal vision and none had any language impairments.

3.3.1.2 Materials and design. Twenty target pictures were selected from the same sources as in Experiment 1. The target pictures in Experiment 2 corresponded to monosyllabic simplex names in Mandarin Chinese (i.e. written using non-decomposable, simplex characters). Each picture was presented with four different types of superimposed monosyllabic distractors: a) semantically related but orthographically and phonologically unrelated (S+O-P-); b) orthographically related but semantically and phonologically unrelated (S-O+P-); c) phonologically related but semantically and orthographically unrelated (S-O-P+); d) semantically, orthographically and phonologically unrelated (S-O-P-).

The distractors in the four conditions, as well as the names of the target pictures, were comparable in terms of word frequency, $F(4, 95) < 1$ (calculated with the log frequency of words in the SUBTLEX-CH database; Cai & Brysbaert, 2010) and visual complexity (number of strokes), $F(4, 95) = 1.421$, $p > .20$. Moreover, two separate online surveys were carried out to ensure the semantically related distractors were not orthographically related to the targets and vice versa. In each survey, 40 native speakers of Mandarin were asked to rate the semantic or orthographic relatedness of word pairs on a 1-7 scale, with the higher score indicating stronger relatedness. Rating scores were first transformed to z-scores per participant, and then submitted to the Friedman test. There were statistically significant differences in the rating scores for orthographic and semantic relatedness among the four conditions, $\chi^2(3) = 71.167$, $p < .001$ and $\chi^2(3) = 67.774$, $p < .001$, respectively. Post-hoc analyses using Wilcoxon Signed-Rank tests were conducted with Bonferroni correction. The results showed respectively that orthographically related stimuli were rated as significantly more orthographically related, and semantically related stimuli were rated as significantly more semantically related compared to the other three conditions, *p*-values < .001. Phonological relatedness was represented by overlapping the segmental information of syllable pairs (e.g., 羊, yang, 'sheep' and 央, yang, 'center'). Twenty other pictures corresponding to monosyllabic names were selected from the same databases to serve as fillers.

The design included two factors: Distractor Type and SOA (-150 ms, -75 ms, 0 ms and 75 ms) as in Experiment 1. In total, there were 16 combinations of the two factors. The 16 conditions were assigned to four groups of participants based on the Latin-square method, with 17 participants per group. In this way, each group of participants was presented with four different combinations of distractor type and SOA, and each saw all the pictures, distractor types and SOAs. In total, each participant received 160 trials (4 blocks by 40 trials).

3.3.1.3 Apparatus and procedure. The apparatus and procedure were the same as in Experiment 1.

3.3.2 Results and discussion

Following the criteria used in Experiment 1, errors (2.61% of all 5,440 data points; including incorrect and disfluent responses) and outliers (0.83%; shorter than 300 ms and longer than 1,300 ms) were excluded from further analysis. Error rates were very low and thus considered not informative enough for further statistical analysis. The naming latencies showed a skewed distribution and were therefore log-transformed. The log-transformed naming latencies (5,253 data points) were submitted to the mixed-effects modelling in R (version 3.1.0; R Core Team, 2014) as the dependent variable.

Table 3.3 The average naming latencies (ms) and percentage errors (in parentheses) for each condition in Experiment 2.

	SOA (ms)			
Distractor type	-150	-75	0	75
Semantically related	657 (.15)	656 (.29)	653 (.26)	588 (.13)
Orthographically related	610 (.17)	621 (.09)	615 (.09)	528 (.06)
Phonologically related	616 (.07)	627 (.11)	627 (.13)	523 (.17)
Unrelated	620 (.09)	632 (.13)	653 (.11)	565 (.11)

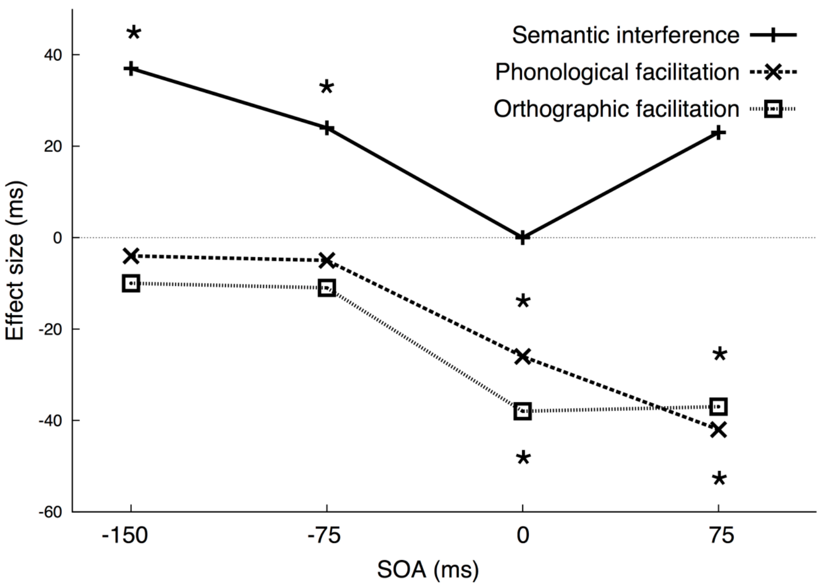


Figure 3.3 The main effects of semantic, orthographic and phonological distractors on picture naming in Experiment 2.

The initial model was built using the 'lmer4' package (Bates et al., 2014) with two fixed factors: distractor type and SOA, the interaction between distractor type and SOA, and one random intercept: target pictures. Since the experiment adopted a between-participants design, the intercept of the particpant was correlated with the fixed factors and thus was not entirely random. The model showed significant interactions between distractor type and SOA, *t*-values > 1.65 (one-tail; based on Zhang et al., 2009; Zhang & Weekes, 2009; Zhao et al., 2012). The data were then divided into four subsets per SOA. Separate models were built with the distractor type as the fixed predictor and random intercept for target picture. The adjusted *p*-values were obtained with the Bonferroni method using the 'multcomp' package (Hothorn, Bretz, & Westfall, 2008).

Table 3.4 The results summary: coefficient estimates, standard errors (SE), *t*-values and *p*-values for the effect of distractor type in each SOA condition in Experiment 2. (significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 "." 0.1 "" 1)

SOA (ms)	Distractor Type	Coefficient Estimate	SE	*t*-Value	*p*-Value
-150	Unrelated	6.41577	0.01832	350.2	
	Semantically related	0.05098	0.01334	3.8	< 0.001***
	Orthographically related	-0.01804	0.01343	-1.3	> 0.05
	Phonologically related	-0.00814	0.01326	-0.6	> 0.05
-75	Unrelated	6.43313	0.01827	352.2	
	Semantically related	0.03496	0.01370	2.6	0.032*
	Orthographically related	-0.02119	0.01351	-1.6	> 0.05
	Phonologically related	-0.00585	0.01352	-0.4	> 0.05
0	Unrelated	6.46080	0.01777	363.0	
	Semantically related	0.00097	0.01431	-0.1	> 0.05
	Orthographically related	-0.05586	0.01462	-3.9	< 0.001***
	Phonologically related	-0.03658	0.01424	-2.6	0.031*

75	Unrelated	6.30905	0.02193	287.64	
	Semantically related	0.02358	0.01919	1.2	> 0.05
	Orthographically related	-0.07703	0.01904	-4.1	< 0.001***
	Phonologically related	-0.07101	0.01911	-3.7	< 0.001***

As shown in Table 3.2, when SOA was -150 ms, there was a significant effect of semantic interference (+37 ms), $p = .0004$. Naming latencies with semantically related distractors were significantly longer than those with semantically unrelated distractors. When SOA was -75 ms, there was again a significant effect of semantic interference (+24 ms), $p = .0321$. The orthographic effect and phonological effect did not reach significance at negative SOAs. These results are in line with the results of Experiment 1.

When SOA was 0 ms, there was a significant effect of orthographic facilitation (-38 ms), $p = .0002$, and a significant effect of phonological facilitation (-26 ms), $p = .0307$. When SOA was 75 ms, there was again significant effects of orthographic facilitation (-37 ms), $p = .0002$ and phonological facilitation (-42 ms), $p = .0007$. The semantic effects did not reach significance at SOAs 0 or 75 ms.

In summary, using solely simplex characters, we did not observe an orthographic effect with negative SOAs, indicating the early orthographic effect shown in Zhang and Weekes (2009) may not be reliably obtained. Instead, both orthographic and phonological effects were found at positive SOAs, replicating results in Zhao et al. (2012). Furthermore, the effect sizes of orthographic and phonological facilitation were also found to be comparable to those in Zhao et al. (2012), i.e. 37 ms and 38 ms after excluding stimuli with phonetic radicals.

In contrast to the results of Experiment 1, at SOA 0 ms, the semantic interference effect did not reach significance in Experiment 2. The discrepancy may be attributed to the difference in distractor frequencies between Experiment 1 and 2. The distractor frequency (calculated by taking the log frequency of words in the SUBTLEX-CH database; Cai & Brysbaert, 2010) is lower in Experiment 1 (mean = 2.49) than in Experiment 2 (mean = 3.64), $p < .0001$. It has been shown that lower-frequency distractors produce stronger interference at the lexical selection stage (Miozzo & Caramazza, 2003). The difference in distractor frequency may also explain the faster average naming latencies and lower error rates in Experiment 2 than in Experiment 1, as due to the less interference during lexical selection in Experiment 2. Note that other possibilities such as differences in stimuli set and/or participant group between the two experiments may also be attributing factors.

Interestingly, when SOA was 0 ms, the orthographic effect ($p = .0002$) was stronger than the phonological effect ($p = .0307$), which is in line with previous findings in English (e.g. Lupker, 1982; Posnansky & Rayner, 1978) and Chinese (Bi et al., 2009). In the present study, the phonological relatedness in the picture-word interference paradigm was presented via orthography, by using Chinese characters. Therefore, it is likely that orthography became available earlier than phonology because phonological relatedness was represented to the speakers via an extra orthography-to-phonology transformation (i.e. phonological information activated after the perception of the characters). Bi and colleagues have tested for independent orthographic and phonological effects as well as their interactions using the picture-word interference paradigm. By using distractors with solely orthographic or phonological relatedness, the grapheme-to-phoneme route (sublexical) may be ruled out and the orthographic relatedness could possibly affect the speech production process via a lexical route (see Bi et al., 2009 for a detailed discussion).

It is worth noting that the distinctive temporal loci of the semantic, orthographic and phonological effects without any overlap in Experiment 2 were similar to the pattern of results found in Experiment 1, which has also been shown for Dutch in Schriefers et al. (1990), where the semantic interference effect was only found at negative SOAs and phonological facilitation at positive SOAs. In both experiments of the present study, the significance of semantic and orthographic effects did not overlap at any SOA. Taken together, these results suggest a two-step model of meaning and form processing during spoken word production for both languages with an alphabetic script like Dutch and languages with a logographic script like Mandarin Chinese. Although additional studies using high temporal resolution measurements such as electrophysiological studies are preferable to settle this debate, the behavioral results of this study do suggest that a general two-step model of speech production that makes distinction between meaning and form processings is sufficient.

3.4 Conclusion

With two behavioral experiments, the present study shows no early orthographic effect, even in a language with a logographic script like Mandarin Chinese where the orthography is characterized by opaque symbol-to-sound mappings. The results run counter to the proposal that orthography affects speech production at an early, conceptual level (Zhang & Weekes, 2009). Rather, the orthographic effects were found at similar temporal loci to the phonological effects, compatible with most speech production models (e.g., Dell & O'Seaghdha, 1992; Levelt et al., 1999a, b; Roelofs, 1992; Roelofs & Meyer, 1998). The results therefore lend further support to a two-step model of speech production in Mandarin Chinese, which distinguishes between meaning and form processings.

Chapter 4

Neural correlates of spoken word production in blocked cyclic naming[5]

[5] A version of this chapter has been submitted for publication as Man Wang, Zeshu Shao, Yiya Chen, & Niels O. Schiller (under review). Neural correlates of spoken word production in blocked cyclic naming.

Abstract

The blocked cyclic naming paradigm has been increasingly employed to investigate the mechanisms underlying spoken word production. In this paradigm, stimuli are presented cyclically in homogeneous and heterogeneous blocks. Semantic homogeneity typically elicits longer naming latencies than heterogeneity; however, it is debated whether competitive lexical selection or incremental learning underlies this effect. The current study investigates spoken word production mechanisms in the blocked cyclic naming paradigm using behavioral and electrophysiological measurements. Both semantic and phonological homogeneity are manipulated to provide evidence that can distinguish between the two accounts. Results show that naming latencies are longer in semantically homogeneous blocks than heterogeneous blocks, but shorter in phonologically homogeneous blocks than heterogeneous blocks. The semantic factor significantly modulates electrophysiological waveforms from 200 ms and the phonological factor from 350 ms after picture presentation. Correlations between naming latency difference and electrophysiological waveform difference are found between semantically homogeneous and heterogeneous blocks in the 200-250 ms time window, suggesting that the semantic blocking effect takes place during lexical selection, and between phonologically homogeneous and heterogeneous blocks in the 500-550 ms time window, suggesting that the phonological facilitation effect reflects strategic preparation. Implications for accounts of word production are discussed.

4.1 Introduction

The blocked cyclic naming paradigm has been increasingly used as a tool to test lexical selection mechanisms during spoken word production. In the blocked cyclic naming paradigm, participants name a small set of pictures either in a homogeneous block (e.g., *apple, peach, pear, orange*) or a heterogeneous block (e.g., *apple, beetle, blouse, duck*; stimuli used in Belke, Meyer, & Damian, 2005) repeatedly in a cyclic manner (Damian, Vigliocco, & Levelt, 2001). In this paradigm, speakers are typically slower in naming pictures in the semantically homogeneous blocks than in the semantically heterogeneous blocks (e.g., Belke, Meyer, & Damian, 2005; Damian et al., 2001; Damian & Als, 2005; Rahman & Melinger, 2009, but see Navarrete, Del Prato, Peressotti, & Mahon, 2014). This is called the semantic blocking effect.

The blocked cyclic naming paradigm is complex in that it involves multiple cognitive components, such as language-specific skills as well as top-down control strategies (e.g., lexical selection, priming, learning, task-representation; Belke et al., 2005; Belke & Stielow, 2013; Oppenheim, Dell, & Schwartz, 2010; Shao, Roelofs, Martin, & Meyer, 2015; see Belke, 2017 for a review). Therefore, it is critical to understand the mechanisms involved in the blocked cyclic naming paradigm in order to use it effectively as a tool to investigate language processing.

One account argues that the underlying mechanism responsible for the semantic blocking effect is competitive lexical selection (Belke et al., 2005; also derived from Howard et al., 2006). Specifically, the previously named picture (e.g. apple) becomes highly active and competes for selection during the subsequent production of a semantically-related target (e.g. peach).

An alternative account argues that competition during lexical selection is not required to produce the semantic blocking effect (Navarrete et al., 2014; also derived from Oppenheim et al., 2010). Instead, such an effect can be explained by an incremental learning mechanism (Oppenheim et al., 2010). This error-based learning mechanism strengthens the connections between the semantic features and to-be-produced words while also weakening the connections between the semantic features and competitors (cf. Spalek, Damian, & Bölte, 2013). More specificially, when the participants name a picture of a mango, the connections between the related semantic features (e.g., fruit, yellow) and the word "mango" are strengthened. Consequently, when the participants name "mango" in the following trials, the naming latcencies will be shorter. By contrast, the connections between the related semantic features and other fruits, such as peach, apple, will be weakened upon naming "mango". Therefore, when naming "peach" or "apple", the naming latencies will be longer. The delay in naming is referred to as "the dark side of incremental learning" (Oppenheim et al., 2010). Navarrete and colleagues (Navarrete et al., 2014) claims that the difference in naming latencies in the blocked cyclic paradigm is caused by the differential priming effects with the underlying incremental learning mechanism. Specifically, in the semantically homogeneous blocks, the connections between the semantic features and target words are weakened for semantically homogeneous words within one cycle, but strengthened for the cyclically repeated target words within a block. By contrast, in the semantically heterogeneous blocks, the connections are always strengthened for the repeated items (i.e. repetition priming). Consequently, naming latencies in the semantically heterogeneous blocks are faster relative to those in the semantically homogeneous blocks where less repetition priming occurs. Navarrete and colleagues (Navarrete et al., 2014) conclude that competitive lexical selection is not required to account for the semantic blocking effect.

Recent studies have made use of electrophysiological and neuroimaging measurements to provide further insights into this debate but have yielded inconsistent findings. By recording the participants' electrophysiological activities in a combination of the picture-word interference and blocked cyclic naming paradigms, Aristei and colleagues (Aristei, Melinger, & Rahman, 2011) found that the semantic blocking effect takes place at around 200 ms after picture presentation. This temporal locus is in line with the locus of lexical selection based on meta-analyses of the temporal and spatial signatures of word production components (Indefrey & Levelt, 2004; Indefrey, 2011). The electrophysiological effect starting around 200 ms after picture presentation is not easily reconciled with Navarette et al.'s (2014) account based on the incremental learning mechanism (Oppenheim et al., 2010). Navarette et al.'s (2014) account predicts less repetition priming in the semantically homogeneous blocks compared to the heterogeneous blocks. Since repetition priming is generally reflected by an attenuated N400 effect (e.g., Rugg, 1985, 1990; see e.g. Misra & Holcomb, 2003 for discussion), less repetition priming should elicit a stronger N400 effect in the semantically homogeneous condition relative to the heterogeneous condition. Furthermore, using neuroimaging and neuropsychological methods, Schnur and colleagues found the semantic blocking effect to be associated with the activities in Broca's area, which corresponds to competition among lexical selection candidates (Schnur et al., 2009; Schnur, Schwartz, Brecher, & Hodgson, 2006). These findings lend support to the competitive lexical selection account. To our knowledge, no supporting electrophysiological evidence has been reported for this account so far.

Alternatively, Janssen and colleagues (Janssen, Hernández-Cabrera, Van der Meij, & Barber, 2015) found a post-retrieval locus of the electrophysiological effect corresponding to the semantic blocking effect represented by longer naming latencies. Janssen et al. (2015) interpreted the "late" effect as a conflict resolution component reflecting an underlying cognitive control mechanism. Therefore, it is still unclear exactly when the semantic blocking effect takes place during spoken word production.

Besides the disagreements on the level of lexical-semantic encoding, another motivation for carrying out the current study is the small number of studies looking into phonological encoding, which is also a critical stage in spoken word production (Indefrey & Levelt, 2004; Indefrey, 2011). The general finding is that when items form a homogeneous block in terms of their onset segments (e.g. *coat*, *cat*, *cook*), naming is facilitated compared to a heterogeneous block, suggesting either facilitation at the word-form encoding stage during speech production or strategic preparation due to high predictability (e.g., Breining et al., 2016; Damian, 2003; Meyer, 1991; Roelofs, 1999; Schnur et al., 2009). However, inhibitory effects have also been observed whether the position of the overlapping segment is not the onset (Breining, Nozari, & Rapp, 2016). Breining and colleagues (2016) suggest a common mechanism responsible for the semantic blocking effect as well as phonological effect; for example, the incremental learning mechanism accounts for the phonological effect in a similar way to the semantic blocking effect in the blocked cyclic naming paradigm.

The present study

The present study aims to contribute to the discussion concerning accounts of encoding in spoken word production by drawing on evidence from the blocked cyclic naming paradigm. With this aim, we probe the semantic blocking effect and the phonological facilitation effect with behavioral and electrophysiological measurements. We hope that by finding the neural correlates of the semantic blocking effect and the phonological facilitation effect, we can better understand the mechanisms underlying spoken word production as reflected by the blocked cyclic naming paradigm.

We present items in semantically homogeneous and heterogeneous blocks, 'homogeneous' meaning that items are congruent in terms of their semantic category and 'heterogeneous' meaning that they are incongruent. Besides semantic congruency, we also investigate phonological congruency: in phonologically homogeneous blocks, items overlap in their onset segment in terms of syllable structure, while in phonologically heterogeneous blocks they do not. Based on the results from previous studies, we expect to observe longer naming latencies in the semantically homogeneous blocks relative to the semantically heterogeneous blocks (e.g., Belke et al., 2005; Belke, 2017; Damian et al., 2001; Damian & Als, 2005; Rahman et al., 2009; but see Navarette et al., 2014), and shorter naming latencies in the phonologically homogeneous blocks relative to the phonologically heterogeneous blocks (e.g., Damian, 2003; Meyer, 1991; Roelofs, 1999; Schnur et al., 2009).

In terms of electrophysiological data outcomes, if competitive lexical selection is involved, we expect to observe a difference in event-related potentials (ERPs) between semantically homogeneous and heterogeneous blocks starting around 200 ms after picture presentation (e.g., Aristei et al., 2011; Indefrey & Levelt, 2004; Indefrey, 2011). Alternatively, a stronger N400 effect is expected in semantically homogeneous blocks relative to the heterogeneous blocks, in line with Navarrete et al.'s (2014) account. Besides this, a correlation in semantic blocks between difference in naming latency and difference in the ERP waveform around 200 ms or 400 ms after picture presentation would lend support to either the competitive lexical selection account or the account by Navarette et al. (2014) based on incremental learning.

If the phonological facilitation effect reflects facilitation at the phonological form encoding stage, we expect to observe ERP differences between phonologically homogeneous and heterogeneous blocks at around 355-400 ms after picture presentation (calculated based on a meta-analysis of the neural correlates of phonological code retrieval and syllabification stages; see Indefrey, 2011 for details). If incremental learning underlies phonological encoding (Breining et al., 2016), a stronger N400 effect is expected in the phonologically homogeneous blocks relative to the heterogeneous blocks. Moreover, a correlation between difference in naming latency and difference in the ERP waveform around 400 ms is expected. Alternatively, if the phonological facilitation effect reflects strategic preparation in performing the naming task, a correlation between the behavioral and electrophysiological data is expected at a later time point near the articulation stage.

4.2 Methods

4.2.1 Participants. Thirty-two native speakers of Mandarin Chinese living in Beijing participated in the study (15 female; mean age = 22.3 years, SD = 3.8 years). They were all right-handed and had normal or corrected-to-normal vision and no history of neurological or language impairment. All participants gave informed consent and received 100 RMB for their participation.

4.2.2 Materials. Thirty-two black-and-white line drawings of common objects were selected from the CRL International Picture Naming Project (Bates et al., 2000) and other standardized picture databases (Snodgrass & Vanderwart, 1980; Zhang & Yang, 2003). Pictures were standardized to 300 by 300 pixels and appeared in the center of the screen as black line drawings on a white background. The target pictures were homogeneous in terms of word length (number of characters, mean = 2.04, SD = .43); and, based on ratings on a 5-point Likert scale, concept familiarity (mean = 4.63, SD = .29), visual complexity (mean = 2.43, SD = .68), subjective word frequency (mean = 3.04, SD = .85), age of acquisition (mean = 5.02, SD = 2.78), and name agreement (the percentage of participants giving the most common name, mean = .81, SD = .12; see Liu, Hao, Li, & Shu, 2011 for details of the norming measurements).

Sixteen of the pictures were selected and combined to create four semantically homogeneous blocks (henceforth S+) with four pictures in each block. The pictures in each block were repeated four times in a cyclic manner. As noted above, the pictures in a semantically homogeneous block belonged to the same semantic category, such as 眼睛 (yan3jing1, 'eye'), 耳朵 (er3duo0, 'ear'), 胳膊 (ge1bo0, 'arm'), 肩膀 (jian1bang3, 'shoulder'). The four blocks contained items belonging to the semantic categories of: animals, clothing, body parts and furniture, respectively. The same sixteen pictures were shuffled and combined to create four semantically heterogeneous blocks (henceforth S-). Twenty native Mandarin speakers who did not participate in the naming experiment were asked to rate semantic relatedness (in term of semantic category) of each set of 4 pictures. The average rating scores were 4.98 (S+) and 1.6 (S-) on a 1-to-5 scale, suggesting the semantically homogeneous blocks were semantically related and the semantically heterogeneous blocks were semantically unrelated.

Another sixteen pictures were selected and combined to create four phonologically homogeneous blocks (henceforth P+) with four pictures in each block. The picture names in a phonologically homogenous block overlapped in their phonological onsets in terms of syllable structure, such as 吉他 (ji2ta1, 'guitar'), 剪刀 (jian3dao1, 'scissors'), 镜子 (jing4zi0, 'mirror'), 金字塔 (jin1zi4ta3, 'pyramid'). There was no overlap in lexical tones. All sixteen pictures were then shuffled and combined to create four phonologically heterogeneous blocks (henceforth P-). The target pictures were considered semantically unrelated based on the average rating scores of semantic relatedness: 1.54 (P+) and 1.32 (P-) on a 1-to-5 scale.

In total, there were sixteen experimental blocks (phonological: 4 homogeneous and 4 heterogeneous and semantic: 4 homogeneous and 4 heterogeneous) resulting in 236 experimental trials. Within each block, each picture was repeated in a pseudo-randomized cyclic manner, i.e. each picture appeared once in each position of the cycle. The sequence of blocks was pseudo-randomized using Mix (Van Casteren & Davis, 2006) so that the same block condition did not appear in two consecutive blocks.

4.2.3 Procedure and apparatus. Participants were seated in front of a monitor at a distance of approximately 50 cm in a soundproof booth. The stimuli were presented using the software E-prime 2.0 and the reaction times (RT) were measured online by a voice-key connected with a PST serial response box. The participants' vocal responses were recorded using the microphone. Incorrect responses were coded manually. Mis-triggered RTs were inspected and corrected manually using the CheckVocal program (Protopapas, 2007).

Before the experiment, the participants were familiarized with the pictures and the names used in the experiment. Each picture was presented once in the center of the screen for 2 s. Following the familiarization, there was a practice session where participants were asked to name the pictures. On each practice trial, a fixation cross appeared in the center of the screen for 500 ms, followed by a jittered blank screen for 500, 600 or 750 ms. Then, the target picture appeared and lasted until the voice-key was triggered or a 2-s limit was exceeded, followed by another blank screen (2 s). Responses that deviated from the names given in the familiarization phase were corrected by the experimenter.

The experimental trial procedure was the same as that of the practice trials. There were four warm-up trials preceding each experimental list, with pictures that were not included as targets. There were self-paced breaks between blocks. The whole experiment lasted about one hour, comprising 30 minutes setting up the electroencephalogram (EEG) equipment and a 30-minute experimental session.

4.2.4 Electroencephalogram recording and data pre-processing. Participants' EEG was recorded simultaneously with 64 Ag/AgCI electrodes using BrainCap (Brain Products GmbH, Germany), following the international 10-20 system. Two EOG electrodes were placed beneath the left eye and at the external canthus of the right eye to record eye movements. On-line recording was referenced to the electrode 'AFz' and the signals were recorded at a sampling rate of 500 Hz. The signals were preprocessed using the Matlab toolbox Fieldtrip (Oostenveld, Fries, Maris, & Schoffelen, 2011). The signals were offline re-referenced to the average of all channels and the data from peripheral electrode sites were excluded to avoid possible muscle activity contamination. The signals of the remaining channels (59) were then band-pass filtered from 0.1 to 30 Hz. ERPs were time-locked to the onset of target pictures and were first segmented from –500 ms to 1000 ms. Artifact rejection was implemented to remove segments with variance values bigger than 1,000 μV^2. Next, an independent component analysis (ICA) was performed in Fieldtrip (code based on a function in EEGLAB; Delorme & Makeig, 2004) to remove the eye-movement artifacts. At most two components per dataset were identified as vertical and horizontal eye movements and removed from the EEG signals for further analysis. The trials were then segmented from -350 ms to 650 ms with a -350 ms to -50 ms pre-stimulus baseline. Trials with amplitudes exceeding ± 100 μV within each trial, or exceeding 5 standard deviations of a participant's mean amplitude of all trials were considered outliers and rejected from the datasets (The cut-off SD value was determined based on visual inspection of five participants' recordings). Datasets from ten participants were excluded due to an insufficient number of remaining trials after artifact rejection and technical problems, leaving twenty-two effective datasets (11 female; mean age = 22.5, SD = 3.8).

4.2.5 Statistical analysis. A total of 2.72% of all data points (5,632) were removed from the behavioral data analysis. This included: (a) incorrect responses; (b) responses with hesitations; (c) voice-key failures (the first three types were considered as errors; the error rate was 2.45% and considered not informative enough for further analysis); (d) outliers (RTs shorter than 200 ms or longer than 1,300 ms; 0.27%). Data (both behavioral and EEG) from the first cycle in each semantic block were also excluded, following a common approach in the blocked cyclic naming paradigm (e.g. Belke et al., 2005).

Altogether 16.36% of all the experimental trials were removed from the ERP data analysis including error trials (2.45%) and segments rejected during artifact rejection (13.91%). There were in total 4,122 trials left for the following analysis. Repeated measures ANOVAs were performed on both behavioral and EEG data.

4.3 Results

4.3.1 Semantic effects. In behavioral data analyses, by-participants and by-items repeated measure ANOVAs were performed with block condition (2 levels: homogeneous vs. heterogeneous) and presentation cycle (3 levels) as two factors. The interaction between the two factors was also included in the model. There was a main effect of semantic relatedness, $F1(1, 21) = 28.315$, $p < .0001$, $\eta^2_p = .574$; $F2(1, 15) = 20.878$, $p < .001$, $\eta^2_p = .582$, demonstrating the semantic blocking effect, i.e. longer RTs in the semantically homogeneous blocks than in the heterogeneous blocks (27 ms; Figure 4.1). There was no significant effect of presentation cycle, $F1(2, 42) = 1.214$, $p = .307$, $\eta^2_p = .055$; $F2(2, 30) = .683$, $p = .513$, $\eta^2_p = .044$. The interaction between block condition and presentation cycle was not significant, $F1(2, 42) = .902$, $p = .413$, $\eta^2_p = .041$; $F2(2, 30) = .583$, $p = .565$, $\eta^2_p = .037$.

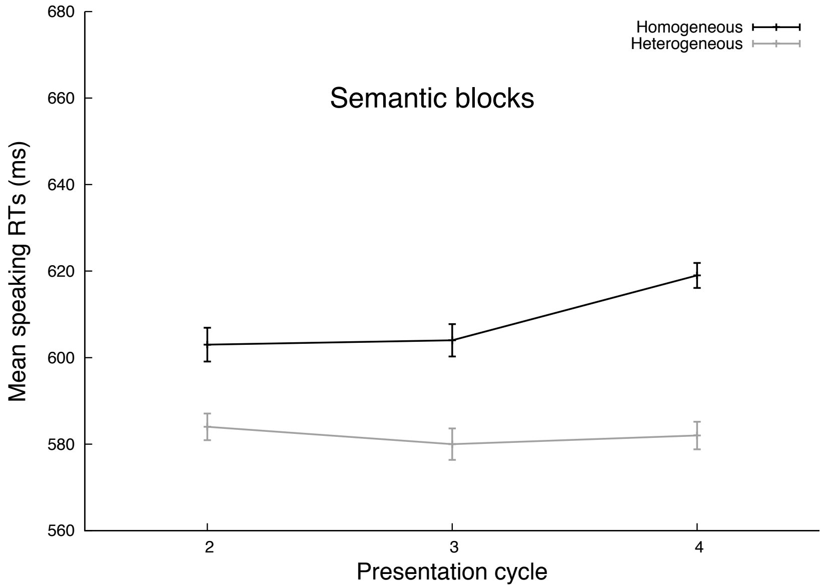


Figure 4.1 The semantic blocking effect in reaction times. Data from the first cycle were excluded (following Belke et al., 2005).

EEG data were also submitted to repeated measures ANOVA, with the mean amplitudes for every consecutive 50 ms time window from 0 ms to 550 ms as the dependent variable and the region of interest (henceforth ROI; 4 levels: left-anterior - F1, F3, F5, FC3, FC5, right-anterior - F2, F4, F6, FC4, FC6, left-posterior - P1, P3, P5, CP3, CP5 and right-posterior - P2, P4, P6, CP4, CP6) and block condition (2 levels: semantically homogeneous versus heterogeneous) as the independent variable (following a similar approach in Costa et al., 2009).

The results showed that in the early time windows (i.e. 0-50 ms, 50-100 ms, 100-150 ms and 150-200 ms), there was only a main effect of ROI, p-values $< .01$, indicating that the mean amplitudes were significantly different between ROIs. Neither the effect of semantic relatedness nor the interaction between ROI and semantic relatedness reached significance.

Between 200-500 ms, there was a main effect of ROI, F-values > 11.0, p-values $< .01$. The interaction between ROI and semantic relatedness was significant, F-values > 4.5, p-values $< .03$. There was a trend of interaction between ROI and semantic relatedness between 500-550 ms, $F = 2.9$, $p = 0.80$. The mean amplitudes per ROI in the semantically homogeneous and heterogeneous conditions were then submitted to pair-wise t-tests, summarized in Figure 4.3. Generally, in the anterior regions, the S- condition elicited more negativities than the S+ condition (see Figure 4.2a). In the posterior regions, the S- condition elicited more positivities than the S+ condition (see Figure 4.2b). The pattern was consistent within 200-550 ms (see Figure 4.2). The detailed effects in each ROI are summarized in Figure 4.3.

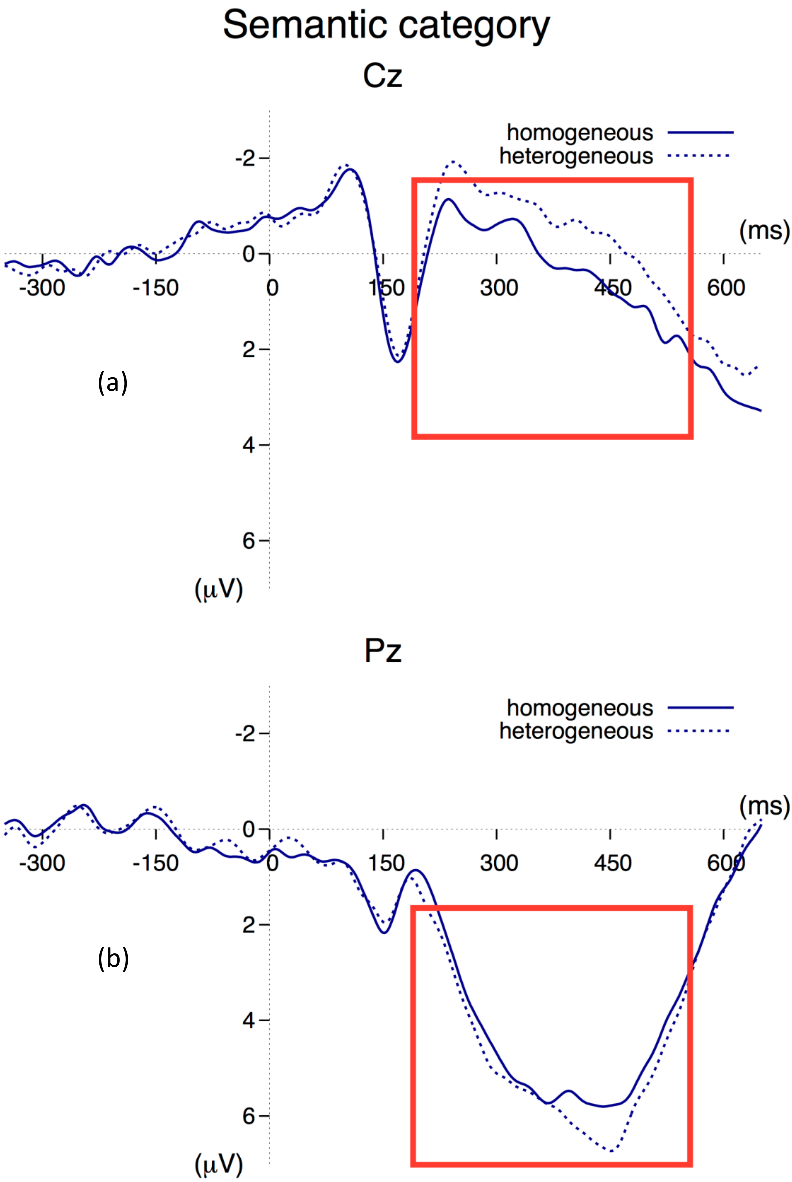


Figure 4.2 The grand average ERPs of the semantically homogeneous (S+) and heterogeneous (S-) conditions. The top graph (a) depicts the ERPs from a representative anterior electrode FC4, with more negativities in the S- than S+

condition. The bottom graph (b) depicts the ERPs from a representative posterior electrode Pz, with more positivities in the S- than the S+ condition.

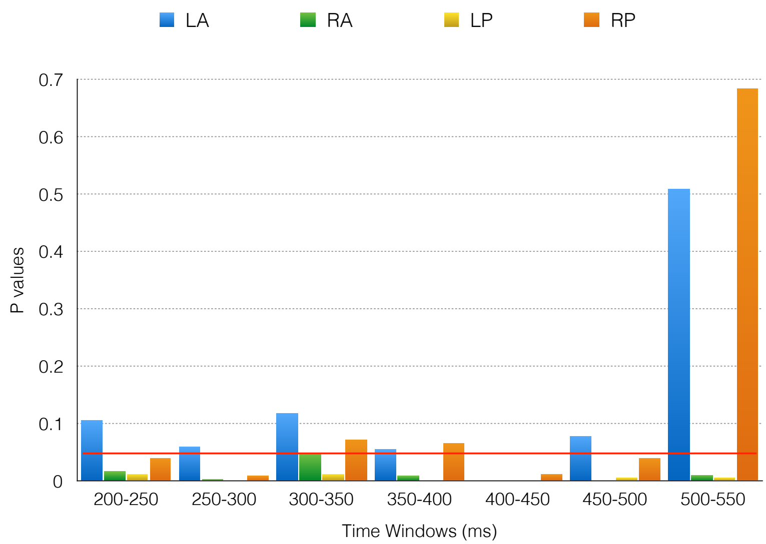


Figure 4.3 The bar graph summarizes the *p*-values resulting from the pairwise t-tests (two-tailed) on the mean amplitudes within each time window per ROI in the semantic blocks. The red line refers to the significance level .05. Four ROIs are represented: left-anterior (blue), right-anterior (green), left-posterior (yellow) and right-posterior (orange).

Besides temporally localizing the semantic blocking effect, we also wished to investigate how ERP effects are related to behavioral outcomes in the blocked cyclic naming paradigm.

If the semantic blocking effect reflects competitive lexical selection, a correlation should be shown around 200 ms after picture presentation, based on the meta-analysis of various electrophysiological studies on word production (Indefrey & Levelt, 2004; Indefrey, 2011). In the present study, within 200-250 ms, there was a main effect of ROI, $F(3, 63) = 15.068$, $p < .001$, $\eta^2_P = .418$. The interaction between ROI and block condition was significant, $F(3, 63) = 5.821$, $p = .007$, $\eta^2_P = .217$. Analysis per ROI revealed significant differences between S+ and S- conditions in the right-anterior, left and right-posterior regions ($p = .016$, .011 and .039, respectively). The mean amplitude difference within this time window per ROI and the RT difference (the heterogeneous condition as baseline) were submitted to the Pearson correlation test. The correlation test showed that the positive correlation in the right-anterior region is marginally significant, $r = .349$, $p = .056$ (see Figure 4.4). Correlations in other regions did not reach significance. These results indicate that the semantic blocking effect observed in RTs correlated with the ERP effect taking place around 200-250 ms after picture presentation, the point at which the lexical selection process takes place according to meta-analyses studies (Indefrey & Levelt, 2004; Indefrey, 2011). The result is in line with the hypothesis that competition during lexical selection underlies the semantic blocking effect. However, it is also possible that what underlies the semantic blocking effect is a sustained process of adjustment, which is most robust after lexical selection (see Belke, 2017), in line with the incremental learning account (Oppenheim et al., 2010; Navarette et al., 2014). Therefore, correlation tests were also performed in the other time windows where significant ERP effects were observed. No correlations were found, p-values > .09.

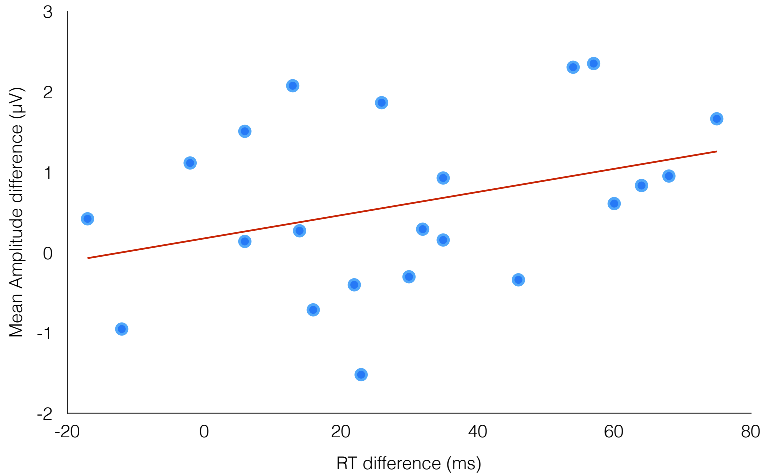


Figure 4.4 This scatterplot depicts the correlation between the behavioral (RT) and electrophysiological (mean amplitude within 200-250 ms in the right-anterior region; the heterogeneous condition as baseline) effects of semantic relatedness in the blocked cyclic naming paradigm.

4.3.2 Phonological effects. In the behavioral data analyses, by-participants and by-items repeated measure ANOVAs were performed with block condition (2 levels: homogeneous vs heterogeneous) and presentation cycle (4 levels) as two factors. The interaction between the two factors was also included in the model. There was a main effect of phonological relatedness, $F1(1, 21) = 11.111$, $p = .003$, $\eta^2_P = .346$; $F2(1, 15) = 11.250$, $p = .004$, $\eta^2_P = .429$, indicating phonological facilitation, with shorter RTs in the phonologically homogeneous blocks than in the heterogeneous blocks (-13 ms). There was also a main effect of presentation cycle, $F1(3, 63) = 50.085$, $p < .0001$, $\eta^2_P = .705$; $F2(3, 45) = 51.976$, $p < .0001$, $\eta^2_P = .776$, indicating that RTs in the later cycles were shorter than in the earlier cycles (see Figure 4.5).

The interaction between block condition and presentation cycle was not significant, $F1(3, 63) = .754$, $p = .524$, $\eta^2_P = .035$; $F2(3, 45) = .893$, $p = .452$, $\eta^2_P = .056$.

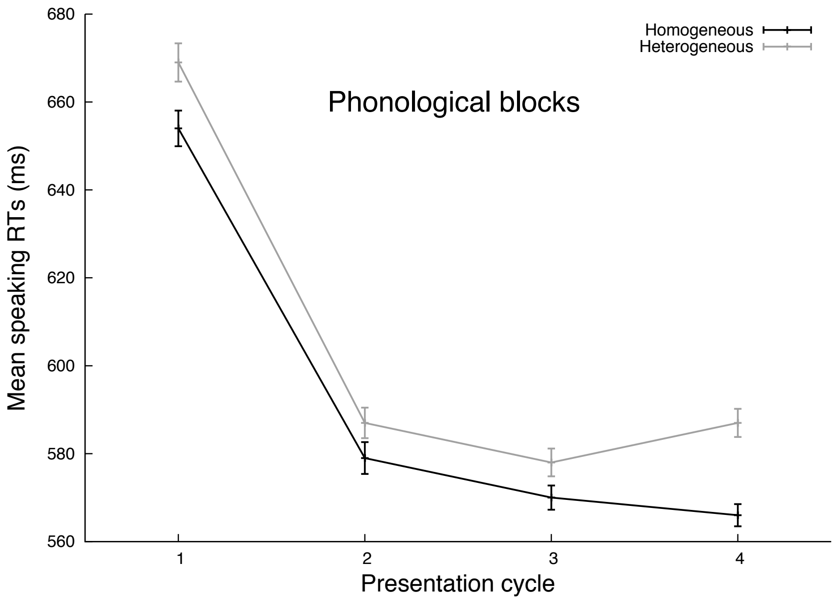


Figure 4.5 The phonological facilitation effect in reaction times across presentation cycles.

In EEG analyses, between 0 and 350 ms, there was only a main effect of ROI, p-values < .01, indicating the mean amplitudes were significantly different between ROIs. Neither the effect of phonological relatedness nor the interaction between ROI and phonological relatedness reached significance.

Between 350-500 ms, there was a main effect of ROI, F values > 13, p-values < .001 and a significant interaction between ROI and phonological relatedness between 350-550 ms, F-values > 3.4, p-values < .05. The mean amplitudes per ROI in the phonologically homogeneous and heterogeneous conditions were then submitted to pair-wise t-tests, summarized in Figure 4.7.

The topographic distribution for phonological effects showed a similar pattern to that of the semantic effects. In the anterior regions, the P- condition elicited more negativities than the P+ condition from 400 to 550 ms (see Figure 4.6a). In the posterior regions, the P- condition elicited more positivities than the P+ condition from 350 to 550 ms (see Figure 4.6b). The detailed effects in each ROI are summarized in Figure 4.7.

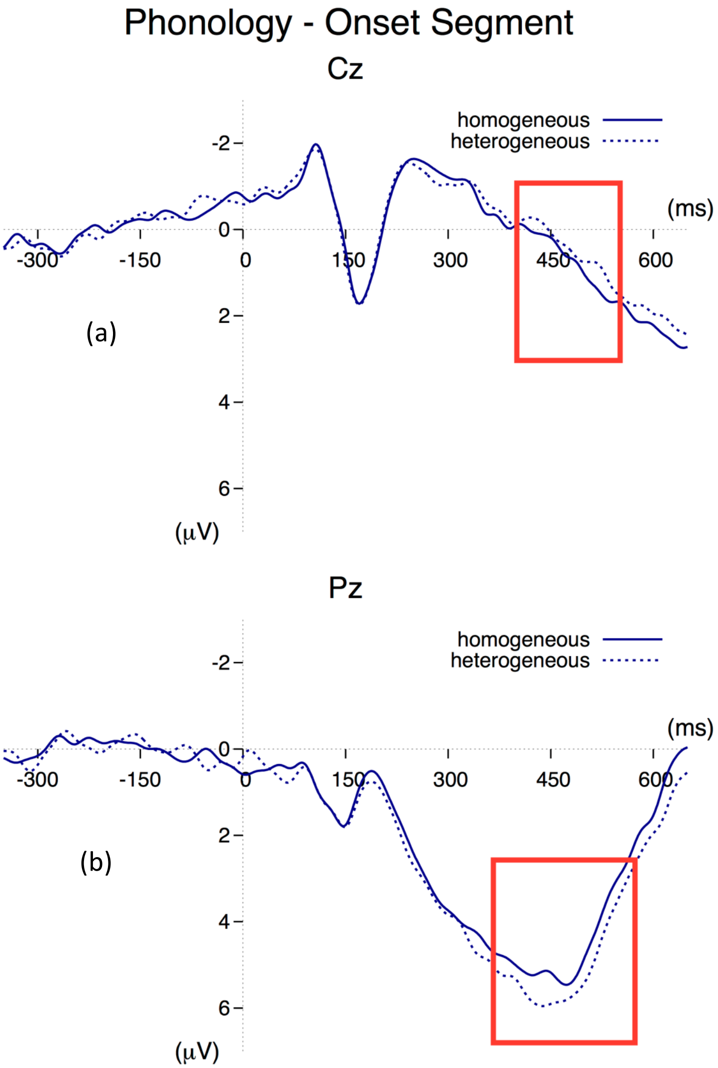


Figure 4.6 The grand average ERPs of the phonologically homogeneous (P+) and heterogeneous (P-) conditions. The top graph (a) depicts the ERPs from a representative anterior electrode FC4, with more negativities in the P- than P+ condition. The bottom graph (b) depicts the ERPs from a representative posterior electrode Pz, with more positivities in the P- than the P+ condition.

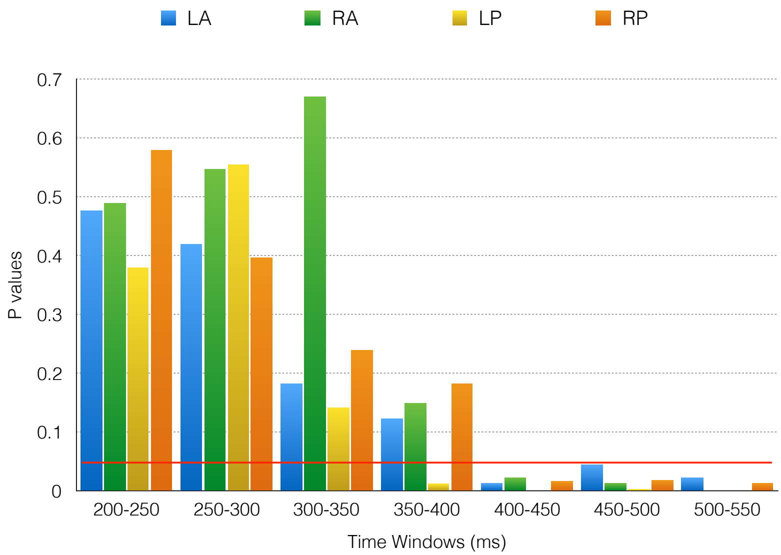


Figure 4.7 The bar graph summarizes the *p*-values resulting from the pairwise t-tests (two-tailed) on the mean amplitudes within each time window per ROI in the phonological blocks. The red line refers to the significance level .05. Four ROIs are represented: left-anterior (blue), right-anterior (green), left-posterior (yellow) and right-posterior (orange).

Correlation tests were performed to assess the relationship between the behavioral and ERP phonological effects. Following the hypotheses outlined in the introduction, we expected the facilitation effect observed in the RTs to localize around 355-400 ms, reflecting facilitation of phonological encoding at the syllabification stage (Breining et al., 2016; Indefrey & Levelt, 2004; Indefrey, 2011), and/or later before articulation reflecting strategic preparation (e.g. Breining et al., 2016). In the time window of 350-400 ms, there was a main effect of ROI, $F(3, 63) = 38.016$, $p < .0001$, $\eta^2_P = .644$ and an interaction between ROI and block condition, $F(3, 63) = 3.429$, $p = .046$, $\eta^2_P = .140$.

Analysis per ROI revealed significant differences between P+ and P- conditions in the left-posterior region, $p = .011$, with the P- condition eliciting more positivities than the P+ condition. The mean amplitude difference within this time window and the RT difference (with the heterogeneous condition as baseline) were submitted to the Pearson correlation test. We expected to see a positive correlation, such that the greater the phonological facilitation, the bigger the difference between the waveforms between the two conditions. However, no correlations were found, suggesting that the phonological facilitation effect observed in the present study may not arise from the phonological encoding process. Correlation tests were also performed in the later time window (500-550 ms) where ERP effects were found. There was a marginally significant negative correlation between the behavioral and ERP differences, $r = -.345$, $p = .058$, indicating that the greater the phonological facilitation, the smaller the ERP effect (see Figure 4.8). This result suggests that the phonological facilitation effect shown in the RT may reflect the strategic preparation in word production. We further discussed the correlations in the discussion.

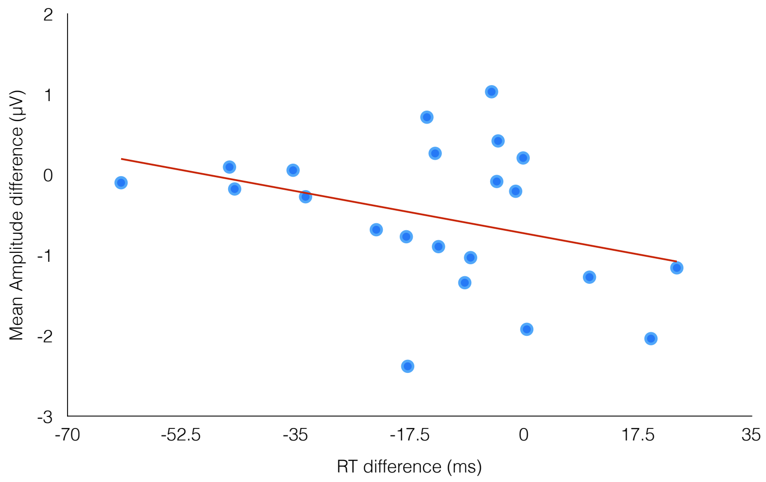


Figure 4.8 The scatterplot depicts the correlation between the behavioral (RT) and electrophysiological (mean amplitude within 500-550 ms in the right-posterior region; with the heterogeneous condition as baseline) effects of phonological relatedness in the blocked cyclic naming paradigm.

4.4 Discussion

Employing behavioral and electrophysiological measurements, we investigated the neural correlates of spoken word production in the blocked cyclic naming paradigm. We observed both the semantic blocking effect and the phonological facilitation effect: reaction times (RTs) in the semantically homogeneous blocks were longer than those in the semantically heterogeneous blocks, in line with previous findings (e.g., Belke et al., 2005; Belke, 2017; Damian et al., 2001; Damian & Als, 2005; Rahman et al., 2009), and shorter RTs were observed in the phonologically homogeneous blocks relative to phonologically heterogeneous blocks, in line with the phonological facilitation effect shown in previous studies (e.g., Damian, 2003; Roelofs, 1999; but see Damian & Dumay, 2009 for an inhibitory effect).

In the electrophysiological data, semantic relatedness modulated the ERP waveforms from about 200 ms and phonological relatedness from about 350 ms after the picture presentation. Correlation analyses showed a significant correlation between behavioral and electrophysiological semantic blocking effects within 200-250 ms, suggesting that the semantic blocking effect takes place during lexical selection (Indefrey & Levelt, 2004; Indefrey, 2011). A correlation between the behavioral and electrophysiological phonological facilitation effects was also observed within 500-550 ms, suggesting that the observed phonological facilitation probably reflects strategic preparation due to high predictability (Breining et al., 2016; Damian, 2003; Meyer, 1991; Roelofs, 1999; Schnur et al., 2009).

In the semantic blocks, significant ERP effects were observed from around 200 to 550 ms after picture presentation. Generally, the semantically heterogeneous condition elicited more negativities in the anterior region and more positivities in the posterior region. The results are thus at odds with the account put forward by Navarrete et al. (2014) based on the incremental learning mechanism. As explained in the introduction, Navarrete et al. (2014) proposed that the semantic blocking effect results from more repetition priming in the semantically heterogeneous blocks than homogeneous blocks. The ERP component N400 is sensitive to both repetition priming and semantic priming, and more priming is associated with an attenuated N400 effect (e.g., Rugg, 1985, 1990; McPherson & Holcomb, 1999; see e.g. Misra & Holcomb, 2003 for discussion and Kutas & Van Petten, 1994 for a review). Therefore, if the incremental learning mechanism underlies the semantic blocking effect, with more priming in the heterogeneous condition (see Navarrete et al., 2014), we would expect the heterogeneous condition to be associated with an attenuated N400 effect. However, the present study yielded attenuated negative effects around 400 ms in the anterior region after picture presentation for the semantically homogeneous condition, contrary to the prediction of the incremental learning account.

The ERP effect in the anterior region bears similarity to the negative effect observed in Cycles 1, 2 and 3 between 250-400 ms in Janssen et al. (2015), with the heterogeneous condition eliciting more negativities. Janssen et al. (2015) interpreted the negative component as reflecting the ease of integrating semantic information in different semantic contexts, with semantic information integration being more difficult in the heterogeneous blocks (Lau et al., 2008). It possibly also reflects the ease of retrieving semantic information from memory (Kutas & Federmeier, 2011; cf. Janssen et al., 2015).

The ERP effect observed in the posterior region has the same polarity as the positive component shown in Cycles 2, 3 and 4 in Janssen et al. (2015) with the heterogeneous condition eliciting more positivities at 500-750 ms, but with an earlier temporal locus in the present study, i.e. 200-550 ms. Janssen et al. (2015) interpreted the positive component as reflecting conflict resolution after lemma retrieval, corresponding to the interference effect observed in the Cycles 2-4 in their study. The average RT in Janssen et al. (2015) is 650 ms, which falls within the time window where the positive component is observed. However, the time window where the positive component is observed in the present study is earlier than the average RT in the semantic blocks (623 ms). Crucially, the chance is small that the effect reflects post-lexical processes especially with the early onset i.e. around 200 ms. Interestingly, we observed a similar component in the posterior region in the phonological blocks, possibly reflecting a task-representation component that is specific to the blocked cyclic paradigm (Belke, 2008; Belke, 2017; Belke & Stielow, 2013). We continue the discussion of the components in the posterior region below.

Correlation tests revealed the relation between the behavioral and electrophysiological data. Correlation tests were performed to assess the relationship between RT difference and ERP waveform difference in the specified time windows and ROIs. A positive correlation was found in the right-anterior region between the semantic blocking effect in the RTs and the mean amplitude difference in the ERP waveforms within 200-250 ms after picture presentation. The process occurring within this time window is assumed to be lexical selection (Indefrey & Levelt, 2004; Indefrey, 2011). The correlation can be easily explained within the competitive lexical selection account (e.g. Levelt et al., 1999), in that the semantically-related items are highly activated in semantically homogeneous blocks and cause competition during lexical selection.

The phonologically heterogeneous condition elicited more negativities in the anterior region and more positivities in the posterior region. In the phonological blocks, significant ERP effects were found from around 400 to 550 ms in the anterior region and from 350 to 550 ms in the posterior region. The topographic distribution is similar for the semantic and phonological effects. The ERP effect in the anterior region resembles the ERP effect associated with phonological priming in the auditory lexical decision task (e.g. Praamstra, Meyer, & Levelt, 1994), with greater phonological mismatch (cf. our phonologically heterogeneous condition) eliciting more negativities. This negative effect also resembles the one found in the semantic blocks, but with a much later onset. This finding is in line with the serial time course proposed for semantic and phonological processes in word production; for instance, using the go/no-go task (e.g. Van Turennout, Hagoort, & Brown, 1997) and the picture-word interference task (Zhu, Damian, & Zhang, 2015). However, the onset of the phonological effect overlaps for at least 150 ms with the time window where the semantic effect is found. This finding indicates that semantic processing precedes phonological processing, but in a cascading or less strictly serial manner.

The ERP effect in the posterior region resembles that observed in the posterior region in the semantic blocks, with a peak around 450 ms after picture presentation. We find it to be close to the P3b component reflecting cognitive workload and/or differences in the probability of pictures seen in homogeneous versus heterogeneous blocks (e.g. Donchin, 1981). The P3b wave "depends on the probability of the task-defined category of stimulus" (Luck, 2005, p. 44). The items in the homogeneous blocks are more predictable within the context of the task than items in the heterogeneous blocks (either semantically or phonologically). Alternatively, this component may correspond to that identified by Belke and colleagues (Belke, 2008; Belke, 2017; Belke & Stielow, 2013), namely, a novel component relating specifically to task representation in the blocked cyclic naming paradigm. The component was proposed based on the observations that when participants have to perform a concurrent digit-retention task, their performances are affected in the blocked cyclic naming task, but not in the continuous naming. Belke and Stielow (2013) point out that in contrast to the continuous naming, the blocked design (homogeneous vs. heterogeneous blocks) means that participants are able to formulate a task-relevant representation and adopt a top-down bias. According to Belke and colleagues' (2013) account, participants can bias the level of activation of words after memorizing the picture set after the first cycle. In the heterogeneous context, the bias-selection mechanism is more efficient because the participants bias only one candidate per semantic category. In the homogeneous context, however, the bias does not help resolve the competition during lexical selection, thus it is more effortful to name pictures in the homogeneous blocks. Ultimately, this account and the probability account are not mutually exclusive. The ERP effects in the posterior region, however, are not easily explained by the account put forward by Navaratte et al. (2014). The reason is that greater priming or ease of adjusting the connections between semantic-lexical features and lexical-segmental features in the heterogeneous blocks would predict an attenuated ERP effect for the heterogeneous condition, rather than the homogeneous condition as observed in the current study.

We found a negative correlation between the RT difference and ERP waveform difference in the right-posterior region for the phonological blocks within 500-550 ms. This time window is rather late in the whole process of speech planning, considering the average RT in the phonological blocks is 599 ms. It is probable that this reflects a process near the articulation stage, based on the meta-analyses of Indefrey and colleagues (Indefrey & Levelt, 2004; Indefrey, 2011). Thus, it is in line with the account that the phonological facilitation effect observed in the RTs probably reflects strategic preparation in the blocked cyclic naming paradigm. We may wonder then why the correlation is negative, i.e. indicating that the stronger the phonological facilitation, the smaller the ERP effect. One possible explanation is that there is high predictability in the phonologically homogeneous blocks, because all the items have the same onset segment. The more the participants adopt strategic preparation, the stronger the phonological facilitation effect is. This strategic preparation then increases the cognitive workload in the phonologically homogeneous condition. Therefore, the ERP waveform in the phonologically homogeneous condition shows more positive deflections and appears closer to that in the phonologically heterogeneous condition.

In summary, in the current study both the semantic blocking effect and phonological facilitation effect were observed in both behavioral and electrophysiological data. Correlation tests between RT differences and ERP differences suggested that the semantic blocking effect takes place during lexical selection, supporting the competitive lexical selection account of spoken word production in the blocked cyclic naming paradigm. Furthermore, the results of correlation analyses indicate that the phonological facilitation effect may reflect strategic preparation due to the high predictability of stimuli in the homogeneous blocks. Distinct but similar ERP effects in the posterior region were observed in both semantic and phonological blocks, with the heterogeneous condition showing more positivities. The positive component is likely to reflect greater cognitive workload, lower predictability of stimuli and may arise due to a task-related top-down selection bias. These results shed light on the neural correlates of blocked cyclic naming and provide novel evidence to further understand the semantic and phonological processes involved in spoken word production.

Acknowledgement

This research was supported by grants from "Talent and Training China-Netherlands" program. We thank Frank Mertz for help with the Matlab script. We thank Elly Dutton for proofreading this manuscript.

Chapter 5

Lexico-syntactic features are activated but not selected in bare noun production: Electrophysiological evidence from overt picture naming[6]

[6] A version of this chapter has been submitted for publication as Man Wang, Yiya Chen, & Niels O. Schiller (submitted). Lexico-syntactic features are activated but not selected in bare noun production: Electrophysiological evidence from overt picture naming.

Abstract

To produce a word, a speaker needs to retrieve the semantic representation of the word and encode the phonological form for articulation. It is not precisely known yet if a word's syntactic features (e.g. number, grammatical gender, etc.) are automatically activated and selected in bare noun production. Using the picture-word interference paradigm, we manipulated the congruency of Mandarin Chinese classifiers (i.e. a lexico-syntactic feature comparable to grammatical gender) between the target picture (e.g. *coat*, classifier-jian4) and the superimposed distractor word (e.g. *luggage*, classifier-jian4 or *rabbit*, classifier-zhi1). The semantic category relatedness was manipulated as well. We measured the participants' naming latencies and their electroencephalogram (EEG). As a result, classifier incongruency elicited a stronger N400 effect in the ERP analyses, suggesting the automatic activation of lexico-syntactic features in bare noun production. However, classifier congruency did not affect naming latencies, suggesting that the lexico-syntactic feature is not selected in bare noun naming when it is irrelevant for production. Implications for word production models are discussed.

5.1 Introduction

Words, together with their semantic, syntactic and phonological properties, are stored in our mental lexicon. When we speak, we access our mental lexicon at an amazingly high speed to select the to-be-produced words to express the meaning in their appropriate phonological forms within the syntactic constraints (Van Turennout, Hagoort, & Brown, 1998). Cognitive language production models predict *when* certain components of a to-be-produced word are activated, selected and encoded, *where* the activation is located in the brain, and *how* the activation flows. Most of these models agree on the main stages involved in word production: (a) conceptualization of the intended message, (b) retrieval of the semantic and grammatical representations of the to-be-produced words (hereafter lemma retrieval), (c) word-form encoding and (d) articulation (e.g., Caramazza, 1997; the spreading-activation model, Dell, 1988, 1990; Dell & O'Seaghdha, 1991, 1992; the WEAVER++ model, Levelt, 1992, 1993; Levelt, Roelofs, & Meyer, 1999a, 1999b; Roelofs, 1992, 1993; Roelofs & Meyer, 1998).

During lemma retrieval, a lemma is activated by the concept and selected for the next stage of phonological form encoding. The word's syntactic features (e.g., number, grammatical gender, etc.) receive activation from the lemma (Figure 5.1). Some syntactic features (e.g. number) may also receive activation from the concepts (e.g. MULTIPLE; Levelt et al., 1999a; see Nickels, Biedermann, Fieder, & Schiller, 2015 for an alternative account). For instance, in English, the *-s* affix needs to be selected for regular plural nouns (e.g. *cats*). In Dutch, the determiner needs to be selected and to agree with the noun on its grammatical gender in noun phrase production (de arm, 'the arm', common gender and het been, 'the leg', neuter gender). Empirical evidence has been reported to support the selection of syntactic features during word and phrase production (e.g., La Heij, Mak, Sander, & Willeboordse, 1998; Schriefers, 1993;

Schriefers & Teruel, 2000; Van Berkum, 1997). Nevertheless, it is debated whether a word's syntactic features (e.g. grammatical gender) are always activated and whether consequently, they are also automatically selected, even when they are irrelevant for specific speech production tasks (e.g., *cat* in English and *been* in Dutch).

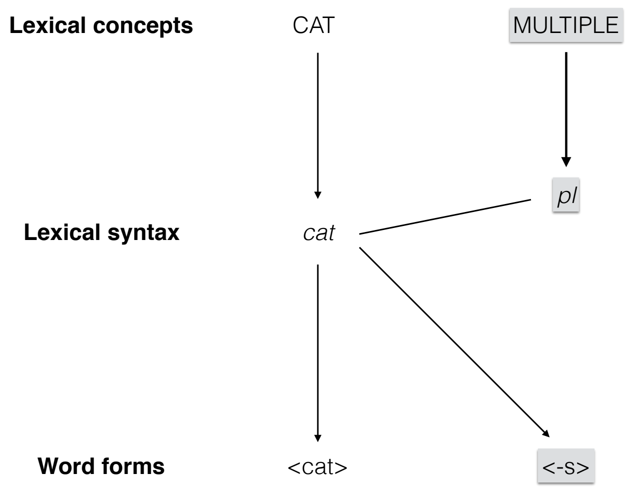


Figure 5.1 The representation of plurals in Levelt et al.'s model (adapted from Levelt et al., 1999a; cf. Nickels et al., 2015, p. 288).

Experimental studies have mostly made use of the picture-word interference paradigm (e.g., Glaser, 1992; see MacLeod, 1991 for a review) to examine the selection of syntactic features in speech production. For example, the selection of grammatical gender in noun phrase production in Dutch and German has been reported (e.g., La Heij et al., 1998; Schriefers, 1993; Schriefers & Teruel, 2000). Specifically, shorter naming latencies were observed when the grammatical gender of the distractor word (e.g., dak, 'roof', neuter gender) was congruent with that of the target picture name (e.g., boek, 'book', neuter gender) than in an incongruent condition (e.g., tafel, 'table', common gender). This has been observed in both article-adjective-noun (e.g., het groene boek, 'the green book') and plain adjective-noun (e.g., groen boek, 'green book') productions. The effect in naming latencies was called the "gender congruency effect" (La Heij et al., 1998; Schriefers, 1993; Schriefers & Teruel, 2000; Van Berkum, 1997), later re-interpreted as determiner congruency effect (e.g., Alario & Caramazza, 2002; Miozzo & Caramazza, 1999; Miozzo, Costa, & Caramazza, 2002; Schiller & Caramazza, 2003, 2006; see Caramazza, Miozzo, Costa, Schiller & Alario, 2001 for a review).

However, no gender or determiner congruency effect was observed in bare noun production in Dutch (e.g., boek, 'book') by La Heij and colleagues (La Heij et al., 1998; see also Starreveld & La Heij, 2004). By contrast, Cubelli and colleagues conducted a series of experiments using the picture-word interference paradigm and reported consistent effects of grammatical gender in bare noun naming in Italian (Cubelli, Lotto, Paolieri, Girelli, & Job, 2005). Therefore, Cubelli and colleagues claim that the selection of grammatical gender is mandatory before accessing the morpho-phonological form of a given noun in word production (Cubelli et al., 2005).

So far, no agreement has been reached upon whether lexico-syntactic features such as grammatical gender are automatically activated and selected in bare noun production. If they are selected as suggested by Cubelli and colleagues (2005), it suggests that speakers select extra information such as task-irrelevant syntactic features in word production. If the lexico-syntactic features are not selected (e.g., La Heij et al, 1998; Starreveld & La Heij, 2004), the theoretical account for the null effect in naming latencies remains unclear. The null effect could be accounted for by speech production models (e.g., Levelt et al., 1999a; Caramazza, 1997) in various ways. One possibility is that the lexico-syntactic features are not activated in bare noun production. The other possibility is that they are always activated but do not affect the retrieval and production of the target word (La Heij et al., 1998).

As discussed in La Heij et al. (1998), even if the lexico-syntactic features are activated, there are still two possible explanations in alternative speech production models. It could be the case that the lexico-syntactic features receive spreading activation from the activated lemma (Levelt et al., 1999a; see Figure 5.1). Since the lexico-syntactic features are activated after the retrieval of the lemma, they will not affect the production speed when irrelevant for production (La Heij et al., 1998). Alternatively, based on the assumptions derived from the model by Caramazza (1997), the syntactic layer (*Lexical syntax* in Figure 5.1) is omitted. The lexico-syntactic information receives activation directly from the semantic representation or the phonological representation. Specifically, the lexico-syntactic features such as word class receive activation from the semantic representation and other features such as gender receive activation from the phonological representation (Caramazza, 1997; cf. La Heij et al., 1998, p. 217).

Therefore, the following questions are empirically open: Are lexico-syntactic features always activated, even in singular bare noun production? If so, where do the lexico-syntactic features receive the activation from (i.e. via spreading activation or direct activation)? Furthermore, are they consequently selected in singular bare noun production?

Note that most studies discussed above have drawn evidence from behavioral studies with reaction time data (but see e.g. Ganushchak, Verdonschot, & Schiller, 2011 for ERP evidence for grammatical gender transfer in Dutch-English bilingualism). Recently, an increasing (though limited) number of electrophysiological studies have investigated the functional characteristics of the language production system, especially the semantic, syntactic and phonological encoding in spoken word production in various picture-naming paradigms (e.g., Indefrey, 2011; Indefrey & Levelt, 2004; see Ganushchak, Christoffels, & Schiller, 2011 for a review). For instance, it has been proposed that the brain engages in lemma retrieval starting 200 ms after stimulus onset (Costa, Strijkers, Martin, & Thierry, 2009; Strijkers & Costa, 2011) and engages in syntactic processing 40 ms before phonological processing during speaking (Van Turennout et al., 1998). Semantic activation has been found to precede phonological encoding during picture naming (Schmitt, Münte, & Kutas, 2000; Van Turennout, Hagoort, & Brown, 1997) as reflected in both the lateralized readiness potentials (LRPs), a derivative of event-related potentials (ERPs), and a response inhibition index, namely the N200. Morphological encoding has been observed around 400 ms after stimulus onset (Koester & Schiller, 2008), in line with the predictions of meta-analytic studies (Indefrey & Levelt, 2004; Indefrey, 2011).

These findings motivated us to seek electrophysiological evidence to tap into the issue of lexico-syntactic feature activation and selection in bare production. Our empirical base for this investigation is bare noun production in Mandarin Chinese. As we will explain below, the nominal classifiers (hereafter classifiers) in Mandarin Chinese provide an interesting as well as important, but hitherto much ignored, test case for the debate.

In Mandarin Chinese, although gender or case is not overtly marked, it is compulsory to use a classifier between a demonstrative and/or numeral and its associated noun. For instance, the common classifier for a piece of upper-body clothing (e.g., *coat*, *shirt*, etc.) is "jian4"[7], and to refer to the noun "da4yi1" (*coat*) in a noun phrase using a numeral or an article, the classifier must occur between the modifier and the noun, i.e. "yi1 jian4 da4yi1" (*one classifier-jian4 coat*) or "zhe4 jian4 da4yi1" (*this classifier-jian4 coat*). Classifier choice is determined by the semantic-syntactic features (e.g., semantic category, number; see Wang, 1973). An example of an object's classifier determined by its semantic category is the contrast between animal names that tend to be used with "zhi1" and clothes names with "jian4". Sometimes classifiers function as the grammatical

[7] As an example, "jian" indicates the phonetic notation of the lexical item, i.e. Pinyin of the word and the number 4 indicates the lexical tone.

marker, comparable to the number morphology in other languages (Cheng & Sybesma, 1999; Cheng & Sybesma, 2005; Doetjes, 1997; Peyraube, 1998).

So far, we have only found two behavioral studies that manipulated classifier congruency as well as semantic relatedness using the picture-word interference paradigm to investigate the role of classifiers in Mandarin Chinese speech production. Conflicting results, however, were reported regarding classifier effects in bare noun naming. Zhang and Liu (2009) found that a classifier-congruent distractor facilitated picture naming even in the bare noun production task where no classifier information was required. However, Wang and colleagues (2006) found contradictory results, and argued that only in noun phrase naming is classifier encoding required, but not in bare noun naming (Wang, Guo, Bui, & Shu, 2006).

In psycholinguistic research, classifier information is considered comparable to grammatical gender information in some respects, as it is directly associated with the lexical item and regarded as a lexical property of nouns. It bears a transparent semantic relationship to the lexical item in some cases, but is arbitrary in others (Tzeng, Chen, & Hung, 1991). Given this similarity, the study of the effect of classifier in noun production is not only necessary but also provides an interesting line of comparison with regard to lexico-syntactic feature encoding between spoken word production in West-Germanic languages (where gender is a prominent feature) and that in East Asian languages (where classification is a prominent feature). In the current study, we used the picture-word interference paradigm and manipulated both semantic category and classifier congruency between target picture name and distractor word. This manipulation provides insights into the classifier choice as a function of semantic classes (e.g. Wu & Bodomo, 2009; but see Cheng & Sybesma, 2005, 2012), which is necessary to tease apart.

We measure both naming latencies and electrophysiological activities. If classifiers are activated as well as selected in bare noun naming, we expect to observe shorter naming latencies on classifier congruent trials than incongruent trials (Zhang & Liu, 2009). As gender disagreement has been reported to elicit a stronger negative effect between 350-500 ms after stimulus presentation (Caffarra, Janssen, & Barber, 2014), we thus expect to observe a reduced N400 effect for the classifier congruent trials, relative to incongruent trials. If classifiers are automatically activated but not selected, we expect to see comparable naming latencies between classifier congruent and incongruent conditions but significant differences between the two conditions in electrophysiological activities. Alternatively, if classifiers are not automatically activated, we expect to see comparable naming latencies and electrophysiological activities between classifier congruent and incongruent conditions. Moreover, we expect to see a general semantic interference effect as reflected in naming latencies, based on previous research using the picture-word interference paradigm (e.g., Glaser & Düngelhoff, 1984; La Heij, 1988; Zhu, Damian, & Zhang, 2015; see Spalek, Damian, & Bölte, 2013 for a review), as well as in the N400 effect due to the semantic integration difficulty (Kutas & Federmeier, 2011; Lau, Phillips, & Poeppel, 2008; Zhu et al., 2015).

5.2 Method

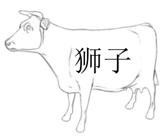
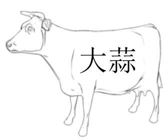
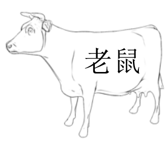
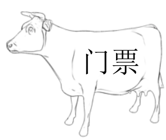
5.2.1 Participants. Thirty-three native Mandarin Chinese speakers (mean age 25 years, SD = 3.05; 19 females) studying in the Netherlands (n = 28) or Beijing, China (n = 5) with comparable second language experience[8] gave informed consent for participation in the experiment. All participants were right-handed, had normal or corrected-to-normal vision, and no history of

[8] A Bartlett test for homogeneity of variance was performed on the behavioral data from the whole dataset, $p > .05$, indicating the homogeneity of the dataset, i.e. the variance does not differ across participant groups recruited in the two locations.

neurological impairments or language disorders. They were paid for their participation.

5.2.2 Materials. Thirty black-and-white line drawings from Severens' picture database (Severens, Van Lommel, Ratinckx, & Hartsuiker, 2005) or similarly drawn, corresponding to monosyllabic (20%), disyllabic (70%) or tri-syllabic (10%) names in Mandarin Chinese served as target pictures. Each picture was presented with four types of distractor words. The distractors were selected based on their congruency with the target picture names regarding two factors – classifier and semantic category (see Table 5.1). The distractors in the four conditions were matched in terms of word frequency, $F(3, 116) = .594$, $p = .620$, number of syllables, $F(3, 116) = 1.790$, $p = .153$, and visual complexity (number of strokes), $F(3, 116) = 1.437$, $p = .236$. Distractors were phonologically and orthographically unrelated to the target pictures.

Table 5.1 An example of a target picture presented with distractor in each condition. Distractors either match or mismatch the classifier (C) or semantic category (S) of target picture name.

	Condition			
Target picture name	C+S+	C+S-	C-S+	C-S-
牛 niu2 classifier-"头"tou2				
distractor	shi1zi0 *lion*	da4suan4 *garlic*	lao3shu3 *rat*	men2piao4 *entrance ticket*
classifier of distractor	"头" tou2	"头" tou2	"只" zhi1	"张" zhang1

5.2.3 Design and Procedure. The experiment adopted a 2-by-2 factorial within-subjects design, with classifier (C) and semantic category (S) as the two factors. Each factor had two levels: congruent (+) versus incongruent (-), resulting in four conditions: C+S+, C+S-, C-S+ and C-S-. On each trial, pictures were presented with a distractor (from one of the four conditions) superimposed on the center of the picture.

All participants saw each of the 30 pictures four times (once for each condition), resulting in 120 trials per participant, which were presented in a pseudo-random order such that the same picture did not occur within ten consecutive trials and no two consecutive trials were from the same condition or with the same corresponding classifier. The pseudo-randomised experimental lists were generated using the Windows program Mix (Van Casteren & Davis, 2006).

The experiment consisted of three sessions: a familiarization session, a practice session and an experimental session. In the familiarization session, each picture was presented once with its name underneath for 2 seconds. Participants were requested to simply view the images and names. In the practice session, each picture was presented once with "XX" superimposed on it and participants were asked to name the pictures with the correct names while ignoring the "XX" on the pictures. Incorrect responses were corrected after the practice session.

In the experimental session, the 120 trials were divided equally into two blocks with a short break in between (length of the break was determined by the participant). On each trial, a fixation point ("+") was presented for 300 ms, followed by a blank screen (200 ms), the target picture with distractor (displayed until the participant initiated a vocal response, with a 2000 ms time-out), followed by another blank screen (500 ms) before the next trial began.

Participants sat in front of a computer in a dimly lit room and were asked to name the pictures using bare nouns as fast and as accurately as possible. Vocal response times were measured by a voice-key and their electroencephalogram (EEG) was recorded simultaneously.

5.2.4 Electroencephalogram recording and data pre-processing. The electroencephalogram (EEG) was recorded using 32 Ag/AgCI electrodes on the standard scalp sites of the extended international 10/20 system. Six flat electrodes were attached above and below the left eye to measure the eye blinks (2), at the external canthus of each eye to record horizontal eye movements (2) and at the mastoids for off-line re-referencing (2).

We used the Matlab toolbox FieldTrip (Oostenveld, Fries, Maris, & Schoffelen, 2011) for the offline processing of the EEG data. The EEG signals were re-referenced to the average of both mastoids and band-pass filtered from 0.1 to 30 Hz. ERPs were time-locked to the onset of the target pictures. Epochs from -200 to 700 ms were computed, including a -200 to 0 ms pre-stimulus baseline. Mean and linear trend were removed from the EEG data using a General Linear Modeling approach prior to resampling the EEG data acquired in two locations (sampled at 512 Hz in the Netherlands and 500 Hz in Beijing) to 256 Hz. We implemented the independent component analysis (ICA) function in FieldTrip (the codes are based on the function of EEGLAB; Delorme & Makeig, 2004) to remove the eye movement artifacts. At most two components per participant were identified as vertical or horizontal eye movements and removed from the EEG signal for further analysis.

Trials with amplitudes exceeding ±100 μV, or a 100 μV difference within a single trial, or exceeding 4 standard deviations of a participant's mean amplitude of all trials were considered as outliers and removed from the analysis. Data from six out of thirty-three participants were excluded from further analysis due to too many artifacts with available segments below 50%

after artifact rejection. The behavioral data from these six participants were excluded from analysis as well.

5.3 Results

5.3.1 Behavioral data. 5.03% of all data points (3,240) were further removed from the behavioral data analysis, comprising: (a) incorrect responses; (b) voice-key failures (the first two types were counted as errors; the error rate was 3.58% and considered not informative enough for further analysis); (c) outliers (i.e. naming latencies exceeding 3 SDs above or below the participant's mean; 1.45%).

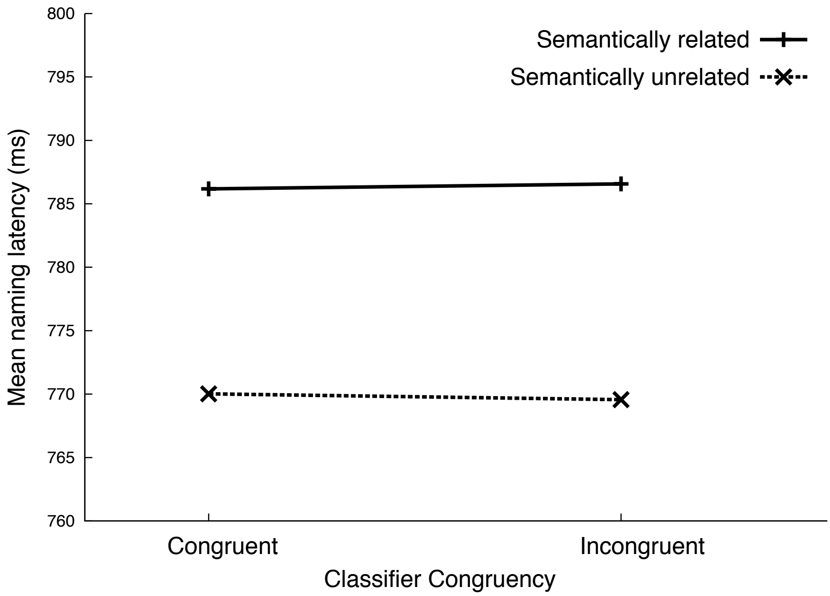


Figure 5.2 There was no significant difference between classifier congruent and incongruent conditions. The naming latencies for semantically related condition were significantly longer than the unrelated condition. There was no interaction between semantic relatedness and classifier congruency.

Repeated measures ANOVAs were performed on the participant means (F1) and item means (F2) with two within-subjects factors: classifier congruency (same classifier vs. different classifiers) and semantic relatedness (same semantic category vs. different semantic categories).

No significant effect of classifier congruency was obtained either in the by-participant analysis, $F1(1, 26) = .000$, $p = .994$, $\eta^2_P = .000$, or in the by-item analysis, $F2(1, 29) = .028$, $p = .867$, $\eta^2_P = .001$, indicating that classifiers are not selected in bare noun naming in Mandarin Chinese. There was a main effect of semantic relatedness in the by-participant analysis, $F1(1, 26) = 14.268$, $p = .001$, $\eta^2_P = .354$ and in the by-item analysis, $F2(1, 29) = 5.041$, $p = .033$. $\eta^2_P = .148$, with longer naming latencies on semantically related trials than semantically unrelated trials (Figure 5.2). The interaction between the two factors was not significant either in the by-participant analysis, $F1(1, 26) = .008$, $p = .928$, $\eta^2_P = .000$, or in the by-item analysis, $F2(1, 29) = .000$, $p = .989$, $\eta^2_P = .000$.

5.3.2 ERP data. 21.02% of all the experimental trials were removed from the ERP data analysis including error trials (3.83%) and segments removed during artifact rejection (17.19%). For each condition, on average, there were 24 remaining segments (1.9 < SDs < 2.3). To avoid possible contamination from eye and muscle movements, data from peripheral electrode sites were not included in the following statistical analysis. Three consecutive time windows (0-275 ms, 275-575 ms, 575-650 ms) were chosen based on previous studies and visual inspection of the data (Figure 5.3; see Zhu et al., 2015 for a similar approach). The mean amplitudes in the above-mentioned time windows across all remaining channels were submitted to repeated measures ANOVA analysis in R (Team, 2014) using the *car* package (Fox & Weisberg, 2011), with classifier congruency (2 levels) and semantic relatedness (2 levels) as two factors.

There was a main effect of classifier congruency, $F(1, 26) = 6.12$, $p = .020$, $\eta^2_P = .191$ and a main effect of semantic relatedness in 275-575 ms, $F(1, 26) = 4.68$, $p = .039$, $\eta^2_P = .153$. The interaction between the two factors was not significant, $p = .50$. No significant effect was found in the other two time windows.

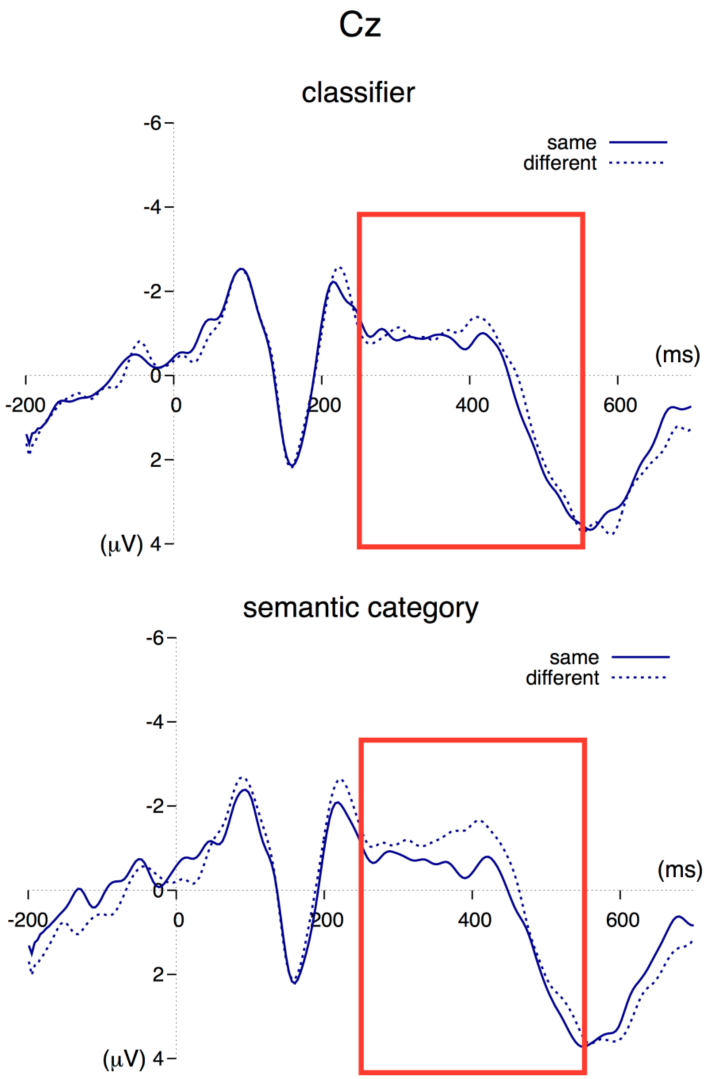


Figure 5.3 *(top)* Grand averages of ERPs in classifier congruent (C+) and incongruent (C-) conditions. Visually, the ERP of C- was more negative ranging from about 275 to 575 ms. *(bottom)* Grand averages of ERPs for semantically related (S+) and unrelated (S-) conditions. Visually, the ERP of S- was more negative ranging from about 275 to 575 ms.

Next, cluster-based permutation tests were performed on each data point (about every 4 ms) to further explore the onset latency and topographic distributions of classifier and semantic effects. Permutation tests (Maris & Oostenveld, 2007) based on *t*-statistics were performed in FieldTrip (Oostenveld et al., 2011) on the participants' mean amplitudes within the time window 275-575 ms where significant semantic and classifier effects were visually observed and statistically confirmed by the ANOVA analysis. This nonparametric randomization test was selected to control for the false alarm rate due to the multiple comparison problem with EEG data. This test first collects the trials into one single set regardless of experimental conditions. A random partition procedure is then performed on the data set 1,000 times and a histogram is constructed of the Monte Carlo approximation of the permutation distribution. The resulting *p*-value reflects the proportion of randomizations that result in a larger test statistic than the observed one. If this p-value is smaller than the critical alpha level of 0.05, then it is concluded that the data between the two experimental conditions are significantly different (see Maris & Oostenveld, 2007 for a detailed description of the method and see e.g. Wang, Bastiaansen, & Yang, 2015 for similar applications of the permutation tests).

Two pairs of comparisons were performed on the amplitudes in the time windows 275-575 ms. We implemented the cluster-based permutation test based on *t*-statistics for all remaining 19 channels (F3, F4, Fz, FC1, FC2, FC5, FC6, Cz, C3, C4, CP1, CP2, CP5, CP6, Pz, P3, P4, PO3, PO4). First the classifier-congruent condition (C+) was compared with the classifier incongruent condition (C-) (both semantically unrelated), and then the semantically-related condition (S+) was compared with the semantically-unrelated condition (S-) (both classifier unrelated). The classifier-congruent and semantically-related condition was omitted (for a similar approach see Zhu et al., 2015).

A significant classifier effect was found from around 370 to 430 ms. The ERP amplitudes were more negative for the incongruent condition than for the congruent condition (Figure 5.4). Similarly, a significant semantic effect was found from around 370 to 430 ms (Figure 5.5). The amplitudes were more negative for the unrelated condition than for the related condition.

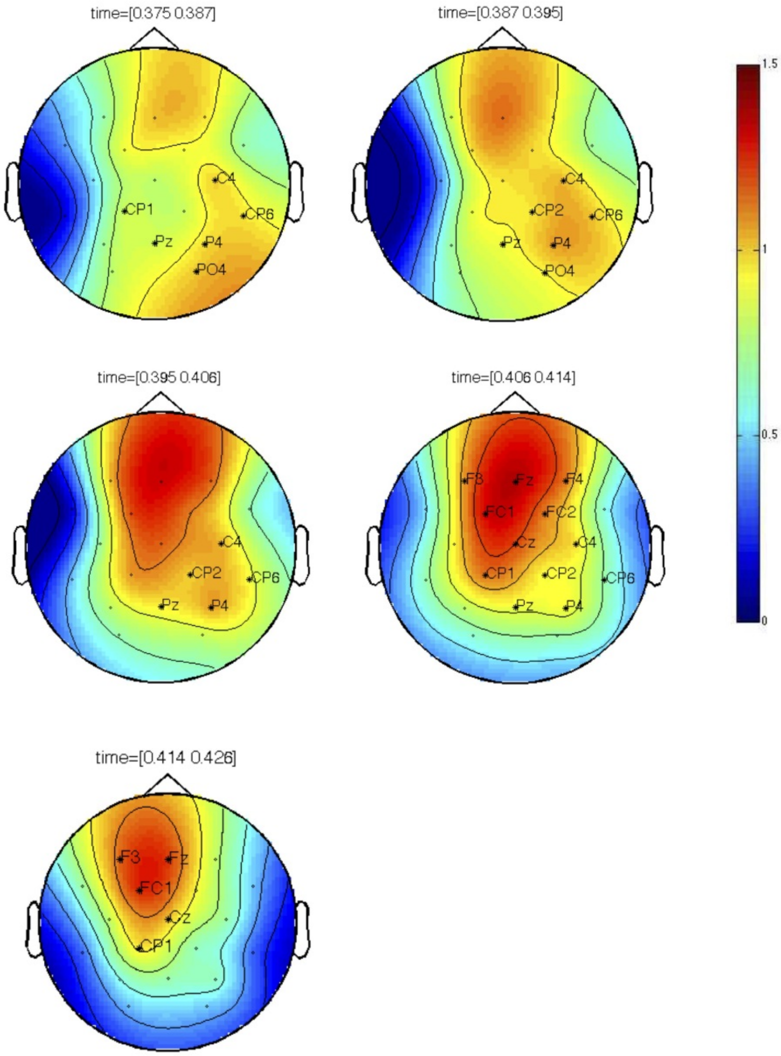


Figure 5.4 A significant positive cluster (C+ minus C-) was found for the classifier effect, ranging from around 370 to 430 ms. Electrodes with significant effects were highlighted with asterisks and channel labels. The topographic distribution was more frontal and right-lateralized relative to that of the semantic effect.

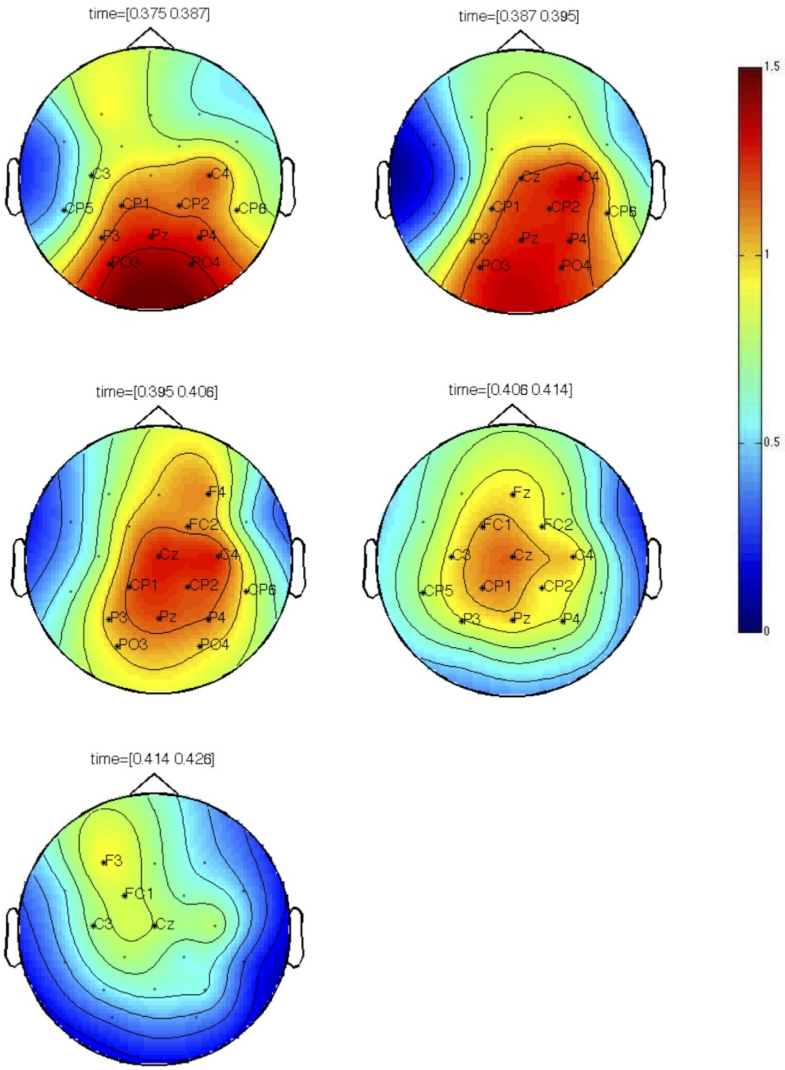


Figure 5.5 A significant positive cluster (S+ minus S-) was found for the semantic effect, ranging from around 370 to 430 ms. Electrodes with significant effects were highlighted with asterisks and channel labels. The topographic plots showed that the semantic effect was most robust in the central-parietal regions.

5.4 Discussion

Using the picture-word interference paradigm, we manipulated the classifier congruency and semantic category congruency between the distractor word and the target picture. By measuring the participants' naming latencies and EEG activities, we investigated if lexico-syntactic features are activated and selected in bare noun production. We will first discuss the semantic effect and then the classifier effect.

The results obtained from manipulating the semantic category were in line with our predictions. The semantic interference effect (e.g., Glaser, 1992; MacLeod, 1991) was revealed by longer naming latencies when pictures were presented with a distractor word from the same semantic category relative to different semantic categories. This is consistent with previous studies (e.g., Glaser & Düngelhoff, 1984; La Heij, 1988; Zhu, Damian, & Zhang, 2015). The semantic interference effect is interpreted as reflecting competition during lexical selection (see, e.g., Levelt et al, 1999a; but see, e.g., Mahon, Costa, Peterson, Vargas, & Caramazza, 2007; see Spalek et al., 2013 for a review).

In the ERP analyses, a larger negative ERP wave was observed for the semantically-unrelated condition compared to the related condition in the time window of 275-575 ms (Figure 5.3). The effect was most robust in the parietal and central regions from about 370 and 430 ms according to a more conservative statistical analysis (Figure 5.5). The ERP modulation by semantic category congruency is consistent with previous studies in Indo-European languages (e.g., Costa et al., 2009; Dell'Acqua et al., 2010; Janssen, Carreiras, & Barber, 2011; Jescheniak, Hahne, & Schriefers, 2003; Jescheniak, Schriefers, Garrett, & Friederici, 2002) and Mandarin Chinese (e.g. Zhu et al. 2015), which also reported greater ERP negativities for the semantically-unrelated condition compared to the related condition. This negative effect at the parietal and central regions and peaking around 400 ms after stimulus presentation resembles a classic N400 effect, elicited by semantic integration difficulty (Kutas & Federmeier, 2011; Lau et al., 2008; Zhu et al., 2015).

No significant classifier effect, however, was observed in the naming latencies of the bare-noun naming task, which is in line with the classifier null effect in bare noun naming reported by Wang et al. (2006) but contradicts the finding of Zhang and Liu (2009). This null effect is similar to that of gender/determiner in Dutch (e.g., La Heij et al., 1998; Starreveld & La Heij, 2004) but different from the grammatical gender effect observed in Italian (Cubelli et al., 2005). Cubelli and colleagues (2005) proposed a two-layer architecture for language production: the lexico-semantic and lexico-syntactic representations. Both layers have to be activated and selected before accessing the phonological form of the target word. To explain the discrepancy between their finding and the null gender effect in Dutch, Cubelli and colleagues (2005) pointed out that only in languages that have a complex morphological structure (e.g. Italian), the selection of grammatical gender is required. Following their suggestion, the null effect of classifier in Mandarin Chinese, a language with a rather simple morphological structure, can be taken as another case for the by-passing of the selection of the lexico-syntactic features in bare noun production.

As discussed in the Introduction, the null effect in naming latencies still leaves open the question of whether the lexico-syntactic features are always activated, even when they are irrelevant for production. Using electroencephalography, we provided fine-grained evidence that supported the automatic activation of the lexico-syntactic features in language production, even in bare noun naming.

A statistically significant effect of classifier incongruency was found between 370-430 ms after the target picture onset (Figure 5.4), albeit in the absence of any significant effect of classifier incongruency in naming latencies. Classifier encoding is not required in bare noun naming, but by manipulating the congruency of classifiers between target pictures and distractors, we observed a reduced N400 effect with classifier congruent compared to incongruent trials. This resembles the negative effect elicited by gender disagreement (Caffarra et al., 2014). The existence of the electrophysiological effect of classifier congruency lends evidence for the automatic activation of classifier features even in bare noun naming.

The remaining question then is how the classifier feature is activated in bare noun naming. There are two possible accounts. Based upon Levelt et al. (1999a)'s model, one possibility is that the classifier receives activation from the activated lemma, as a lexico-syntactic feature. Since this process happens after the lemma retrieval, we then would not expect the activation to affect the naming latency. Alternatively, based upon the Caramazza's (1997) model, the other possibility is that the classifier, as a lexico-syntactic feature, receives activation directly from semantic representations or phonological representations. We know that classifiers in Mandarin Chinese can be independent from both the semantic representation and the phonological representation. For instance, native speakers of Mandarin Chinese acquire the classifier-noun combinations around four and five years old (e.g., Erbaugh, 1986; Fang, 1985) and 'there is no transparent or unequivocal mapping between conceptual properties and classifiers' (cf. Bi, Yu, Geng, & Alario, 2010, p. 103). As a consequence, the correct classifier-noun combinations have to be memorized. Therefore, it is more likely that it is the activated lemma that spreads activation to the classifier feature, rather than activation directly from semantic or phonological representations.

The topographic and temporal demonstrations of the semantic and classifier effects lend further support to this lemma activation account. We observed a more robust semantic effect before 400 ms while the classifier effect was more robust after 400 ms (Figures 5.4 and 5.5). Moreover, the effect appeared to be less robust based on the grand averages of ERPs compared to the semantic effect (Figure 5.3). Consistent with what is shown with the grand averages of ERPs, the effect was shown in a smaller region than the range of electrodes displaying a significant semantic effect (Figures 5.4 and 5.5). Conjointly, these results support the possibility that the classifier feature receives activation from the target lemma.

In Figure 5.6, extending the speech production model from Levelt et al.'s (1999a), we show that for the lexical concept COW, the consequently activated target lemma (e.g. 牛, niu2, 'cow') automatically spreads the activation to the classifier feature (e.g. *classifier* 头, tou2, 'head') of this target lemma via Link A. When we have a distractor word (e.g. 门票, men2piao4, 'entrance ticket'), which also activates its lemma and automatically its classifier (e.g. *classifier* 张, zhang1, 'piece') that differs from that of the target (头, tou2, 'head'), it elicits a stronger N400 effect, relative to the condition where a distractor (e.g. 大蒜, da4suan4, 'garlic') has the same classifier as that of the target (e.g. *classifier* 头, tou2, 'head'). However, in bare noun naming where the classifier information is not required for production, the incongruency between different classifier features does not affect the naming latencies.

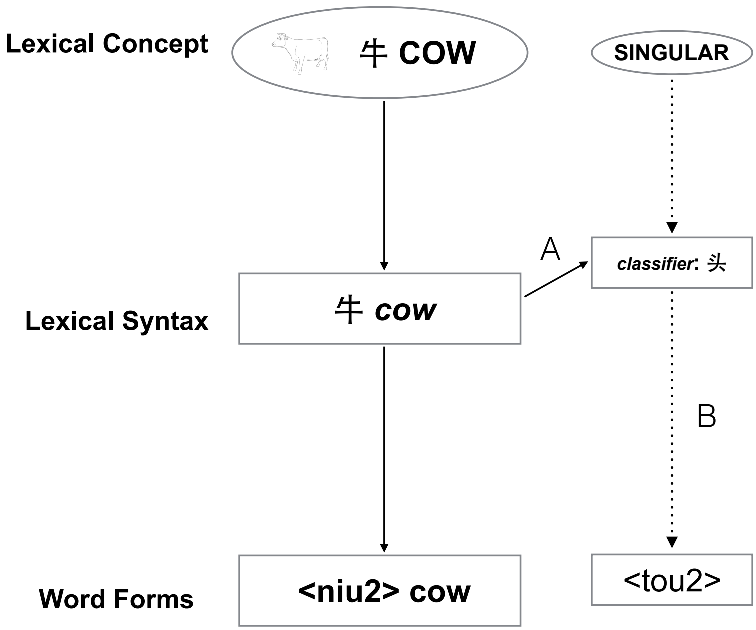


Figure 5.6 The automatic activation of the lexico-syntactic representation of classifiers in word production of Mandarin, adapted from Levelt et al. (1999a). The phonological form encoding of classifiers is not necessary in bare noun naming so Link B is only present when the production of classifier is required. Other lexico-syntactic features such as number and case that require more on-line processing rather than retrieval from long-term memory are not included in this framework.

To conclude, our behavioral and electrophysiological results jointly suggest that the Mandarin classifier feature is automatically activated by its associated target lemma but it is not selected in bare noun naming. Future research can be beneficial to further investigate to what extent automatic activation of lexico-syntactic features is language universal.

Acknowledgments

This research was supported by grants from the "Talent and Training China-Netherlands" program. We thank Gareth O'Neill, Jenny Doetjes and Niels Janssen for their comments on this study. We thank Frank Mertz for help with comprising the Matlab script.

Chapter 6

General Discussion

This dissertation investigates the speech production of Mandarin Chinese from a psycholinguistic approach. Why is it interesting to investigate Mandarin Chinese speech production? From a theoretical point of view, current psycholinguistic models of speech production have been mainly based on evidence from West Germanic languages, where orthographic and phonological forms follow a certain mapping captured in grapheme-to-phoneme conversion (GPC) rules. By contrast, in languages with a logographic script such as Mandarin Chinese, GPC is more opaque, which may result in a (different) role for orthography in speech production. Previous research on the speech production of languages with a logographic script has also provided empirical evidence suggesting the need of modifications to the current speech production models (e.g. Qu, Damian, & Li, 2016; Verdonschot, 2011; Zhang, Chen, & Weekes, 2009; Zhang & Weekes, 2009).

This dissertation provided direct evidence, first with reaction time measurements, for the involvement of orthography in speech production in Mandarin Chinese (Chapter 2) and that the orthographic effect on speech production was rather independent. That is, the orthographic representation of a lexical item exerted its effect without interacting with its semantic or phonological representations (Chapter 3). The following chapter then provided electrophysiological evidence supporting relatively early semantic processing and relatively late phonological form encoding in Mandarin Chinese (Chapter 4) as well as electrophysiological evidence supporting the automatic activation of lexico-syntactic features in speech production of Mandarin Chinese (Chapter 5).

Chapter 1 introduced the current psycholinguistic models of speech production. Most models agree that to overtly produce a word, speakers go through several stages: conceptual preparation, lemma retrieval, word-form encoding and articulation (e.g., Caramazza, 1997; Dell & Seaghdha, 1991, 1992; the WEAVER++ model, Levelt et al., 1999a, b; Roelofs, 1992; Roelofs & Meyer, 1998). At the word-form encoding stage, word form usually refers to the phonological form of the word. Note that the Independent Network theory does specifically recognize an orthographic representation and a phonological representation of the lexical item, but only hypothesizes a role of orthography in written word production (e.g. Caramazza, 1997; Rapp & Caramazza, 2002).

Subsequently, this dissertation pointed out that in languages with a logographic script like Mandarin Chinese, the orthographic representation of a lexical item - Chinese characters had a critical role in distinguishing homophones and might therefore be involved in speech production. Furthermore, the speech production mechanisms of Mandarin Chinese might differ from the predictions of current speech production models.

As the first experimental chapter, **Chapter 2** directly tapped into the question whether orthography was involved in speech production of Mandarin Chinese. No consensus has been reached in terms of the involvement of orthography in speech production. Empirical evidence was reported to suggest the mandatory activation of orthography in speech production in English (Damian & Bowers, 2003) in the form-preparation paradigm (Meyer, 1990, 1991). More specifically, inconsistent spelling (e.g. 'giant', 'jewel', 'joker') in a phonologically homogeneous context disrupted the form-preparation effect. The authors (Damian & Bowers, 2003) also conducted a post hoc analysis on previous studies in Dutch (Meyer, 1990, 1991) but did not find a similar disruptive effect caused by the orthographic inconsistency. Similarly, in a later study using visually masked primes to test reading aloud in Dutch, the orthographically related primes (e.g. 'cement', *concrete*) did not speed up the reading responses of the targets (e.g. 'congres', *congress*) (Schiller, 2007). Moreover, the mandatory involvement of orthography was not observed in French (Alario, Perre, Castel, & Ziegler, 2007) or Chinese (Chen, Chen, & Dell, 2002).

One possible explanation for the discrepancy is that the involvement of orthography may be task-dependent. For instance, orthographic inconsistency showed an inhibitory effect in a word-reading task in the form-preparation paradigm but not in picture naming, word generation or associative naming. This task-dependent characteristic is consistent in Dutch (Roelofs, 2006) and Chinese (Bi, Wei, Janssen, & Han, 2009). These findings seem to suggest that only in tasks where the orthographic information is highly relevant, there may be the involvement of orthography in speech production. Another possibility is that the discrepancy may be attributed to the cross-linguistic differences. As discussed in Damian and Bowers (2003), compared to Dutch, in English the mapping between orthography and phonology is more opaque, which may result in the involvement of orthography in speech production in English.

Aiming to resolve the discrepancies, in Chapter 2, we re-investigated the role of orthography in Mandarin Chinese using an adapted blocked cyclic naming paradigm. In this paradigm, participants were asked to overtly name pictures that were presented repeatedly in semantically homogeneous, phonologically homogeneous, or heterogeneous blocks. On each trial, a written Chinese character that was either orthographically related or unrelated to the target was briefly presented (for 75 ms) before the target picture. We measured participants' speech onset latencies. Consistent with previous research, an inhibitory semantic blocking effect and a facilitative phonological blocking effect were found. More importantly, we observed that the orthographically related characters facilitated picture naming in both the semantic and phonological blocks. In addition, the orthographic priming effect was independent of both the semantic and the phonological effects. These findings suggested that orthography contributes to speaking in a picture naming task, lending further support to the presence of orthographic priming in spoken word production, even in a language with a logographic script like Chinese.

The contribution of orthography to speech production in Mandarin Chinese lent support to the suggestion that in a language with relatively opaque mapping between orthography and phonology, orthography was involved in speech production (Damian & Bowers, 2003). As for the claim that orthography was only involved when highly relevant for production (e.g. in reading tasks; Roelofs, 2006), we offered extra empirical evidence for future discussions. In the adapted blocked cyclic naming paradigm, the Chinese characters were very briefly presented and the participants barely had time to consciously process the characters. Still, an orthographic priming effect was demonstrated. This finding contrasted with the null effect of orthography in picture naming in Chinese (Bi et al., 2009), however, the contrastive results could be attributed to various reasons (e.g. stimuli sets, participant groups, experimental task).

In Chapter 2, we found an orthographic facilitation effect, indicating that the activation of an orthographic representation could facilitate lexical access in spoken word production. The effect was present from the first cycle in the blocked cyclic naming paradigm with orthographic priming, and thus could not have originated from a learning phase (see Alario et al., 2007).

Chapter 3 investigated when and how orthography was involved during speech production. In previous research, the orthographic effect was observed at a similar stage to the semantic effect without the co-occurrence of any phonological effect. It was then suggested that orthography affected speech production via a lexico-semantic pathway (Zhang et al., 2009; Zhang & Weekes, 2009; Figure 6.1). The critical evidence that supported this claim was that the orthographic effect was observed at negative SOAs but this observation was not replicated in a later study (Zhao, La Heij, & Schiller, 2012). This chapter attempted to replicate it but did not observe any orthographic effect at negative SOAs in Experiments 1 or 2, suggesting that it was unlikely that orthography affected speech production of Mandarin Chinese via a lexico-semantic pathway.

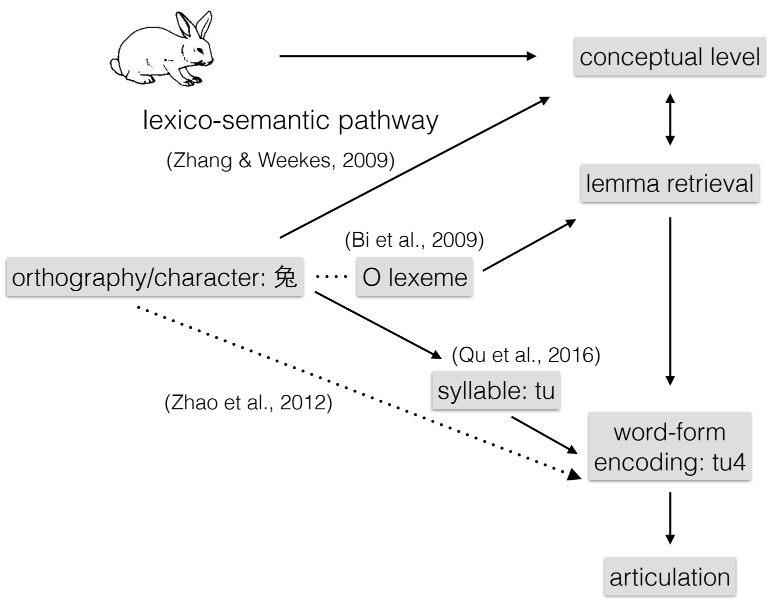


Figure 6.1 An overview of predications on the orthographic effect on speech production of Mandarin Chinese.

In Experiment 2 of Chapter 3, we took a step further and used simplex characters only so as to clearly dissociate orthography from the semantic representation and phonological representation. Consistent with the finding in Zhao et al. (2012), the orthographic effect was observed with the co-occurrence of the phonological effect, subsequent to the semantic effect. In previous research, Bi and colleagues (Bi, Xu, & Caramazza, 2009) elaborately discussed the possible routes of the orthographic effect on speech production. The authors suggested that pure orthographic relatedness (i.e. semantically and phonologically unrelated) facilitated speech production at the lexical level (Bi et al., 2009; Figure 6.1). More specifically, the orthographically related distractors activated the orthographic neighbors, including the orthographic representation of the target and activation spread to the target lemma (Bi et al., 2009). If this was the case, we should have observed that the orthographic effect arose at a similar stage to the semantic effect. Such a pattern, however, was not observed in Experiments 1 or 2.

Alternatively, Zhao and colleagues (Zhao et al., 2012) claimed that the orthographic relatedness might affect speech production at a similar stage to the phonological relatedness, i.e. the word-form encoding stage. Nevertheless, in speech production, orthographic word form encoding is not necessary. Therefore, the only way for orthography to affect the word-form encoding stage is to facilitate the phonological form retrieval and encoding. We made use of the simplex characters, i.e. characters without phonetic radical, so that the GPC route was ruled out as a possible pathway. This suggests that orthographic relatedness may affect another sub-lexical level, i.e. the character-to-syllable correspondence (Qu et al., 2016; Figure 6.1). More specifically, for a target (e.g. 兔, tu4, 'rabbit'), the orthographically related distractor (e.g. 免, mian3, 'exemption') activated its orthographic neighbors (e.g. 兔, tu4, 'rabbit'), which, consequently, activated character activated its syllable (tu4) and facilitated the speech production of the target.

Besides drawing evidence from behavioral data, in recent decades, researchers have increasingly used electrophysiological measurements to investigate the underlying mechanisms of speech production (Christoffels, Firk,

& Schiller, 2007; Koester & Schiller, 2008; see Ganushchak, Christoffels, & Schiller, 2011 for a review). With the high temporal resolution of electrophysiological measurements, Chapters 4 and 5 tapped into the time course and the neural correlates of speech production of Mandarin Chinese.

Chapter 4 investigated the neural correlates of semantic and phonological processing in speech production of Mandarin Chinese. Firstly, consistent with the findings in Chapter 2 and previous research, longer naming latencies were shown in semantically homogeneous blocks and shorter naming latencies in phonologically homogeneous blocks, relative to the heterogeneous blocks. Then, in the electrophysiological data, it was shown that the semantic factor significantly modulated electrophysiological waveforms from 200 ms and the phonological factor from 350 ms after picture presentation. The results were consistent with the estimation of meta-analyses on the neural correlates of speech production (Indefrey & Levelt, 2004; Indefrey, 2011; Strijkers, Costa, & Thierry, 2010) and studies using the go/no-go task (e.g. Van Turennout, Hagoort, & Brown, 1997) and the picture-word interference task (Zhu, Damian, & Zhang, 2015). This suggested that the speech production of Mandarin Chinese also involved an earlier semantic processing and a later phonological processing and the temporal loci of these two stages were in line with those of the estimation of speech production in general.

The previous chapters tested the semantic, orthographic and phonological processing during speech production of Mandarin Chinese. **Chapter 5** tapped into a more specific detail in the process of speech production; that is, whether a word's syntactic features (e.g. number, grammatical gender, etc.) were automatically activated and selected in bare noun production. Previous research has shown that the lexico-syntactic features are activated and selected in noun phrase production when these features are necessary for production (e.g., de arm, 'the arm', *common gender*, 'het been', 'the leg', *neuter gender*; see Caramazza, Miozzo, Costa, Schiller & Alario, 2001 for a review). However, it has been debated if the lexico-syntactic features are activated and selected in bare noun production when these features are irrelevant for production (e.g., 'arm', 'been'; see, La Heij et al., 1998; Starreveld & La Heij, 2004 for a null effect of grammatical gender in bare noun production in Dutch; Cubelli, Lotto, Paolieri, Girelli, & Job, 2005 for an effect of grammatical gender in bare noun production in Italian; Tsegaye, Mous, & Schiller, 2014 for an effect of plural gender and masculine/feminine gender noun productions in Konso). In Mandarin Chinese, although gender or case is not overtly marked, it is compulsory to use a classifier between a demonstrative and/or numeral and its associated noun. In psycholinguistic research, classifier information is considered comparable to grammatical gender information (Tzeng, Chen, & Hung, 1991).

Using the picture-word interference paradigm, we manipulated the congruency of Mandarin Chinese classifiers between the target picture (e.g. 'coat', *classifier-jian4*) and the superimposed distractor word (e.g. 'luggage', *classifier-jian4* or 'rabbit', *classifier-zhi1*). We measured the participants' naming latencies and their electroencephalogram (EEG). As a result, classifier incongruency elicited a stronger N400 effect in the ERP analyses, suggesting the automatic *activation* of lexico-syntactic features in bare noun production. However, classifier congruency did not affect naming latencies, suggesting that the lexico-syntactic feature was not *selected* in bare noun naming when it was irrelevant for production. The null effect of classifier congruency in naming latencies was in line with the results in Wang, Guo, Bi and Shu (2006) for Chinese and Dutch (La Heij et al., 1998; Starreveld & La Heij, 2004) but contradicted the results in Zhang and Liu (2009) in Chinese and Italian (Cubelli et al., 2005). It is possible that for speech production in languages with relatively simple morphological structures, the selection at the lexico-syntactic layer is not necessary (see Cubelli et al., 2006 for a detailed account of a two-layer architecture for language production). Moreover, the automatic activation of classifier information may be attributed to the fact that Mandarin speakers acquire and memorize the classifier-noun combination at very young ages and the classifier feature receives activations spread from the activated lemma.

In short, this study of classifier effects provided insights to the comparison with regard to lexico-syntactic feature encoding between spoken word production in West-Germanic languages (where gender is a prominent feature) and that in East Asian languages (where classification is a prominent feature).

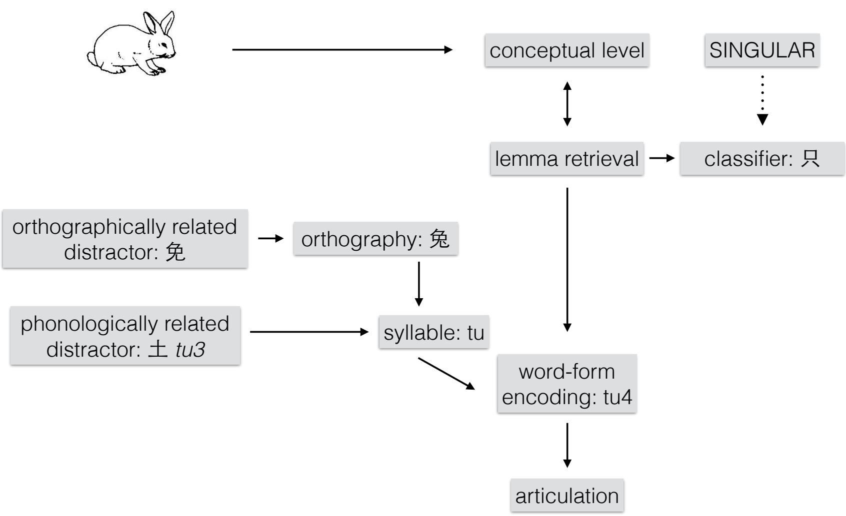


Figure 6.2 The speech production of Mandarin Chinese in the context of orthographically, phonologically or classifier related distractors.

Conclusion and implications for future research

In summary, this dissertation investigated the speech production processes and mechanisms in Mandarin Chinese from a psycholinguistic perspective. The characteristic opaque grapheme-to-phoneme correspondence in Chinese provides an interesting test case for speech production, especially with regard to the separate roles of orthography and phonology. Results of the experiments reported in this dissertation show that orthography contributes to speech production, probably by activating its orthographic neighbor and then the corresponding target syllable. In addition, pure orthographic relatedness can affect speech production independently without interacting with semantic or phonological representations. Moreover, this dissertation used electrophysiological measurement to investigate the fine-grained time course of speech production in Mandarin Chinese. It was shown that the semantic factors modulated the electrophysiological signals from 200 ms and the phonological factor from 350 ms after stimulus presentation. It was also shown that the lexico-syntactic feature (Chinese classifier) was automatically activated in speech production even when it was not necessary for production.

This dissertation tapped into the semantic, orthographic and phonological effects in speech production in Mandarin Chinese in the framework of current psycholinguistic models of speech production. The findings in this dissertation not only contribute to the understanding of the underlying neuropsychological mechanisms of speech production in Mandarin Chinese, but also provide insights into the understanding of the accountability of current models of speech production that are mostly based on evidence from West Germanic languages.

For future studies, it would be interesting to look into the proximate unit of phonological encoding in speech production of Mandarin Chinese. It has been debated that the proximate unit is the syllable (Chen et al., 2002; O'Seaghdha, Chen, & Chen, 2010), the phonemic segment (Qu, Damian, & Kazanina, 2012), or a sub-syllabic unit (Verdonschot et al., 2015). While investigating the phonological encoding in this dissertation, it was shown that the sub-syllabic overlap modulated brain signals from 350 ms after stimulus presentation (Chapter 4). The existence of this effect may be attributed to the phenomenon that young generations of speakers of Mandarin Chinese, our participant population, tend to type *pinyin* (the phonetic notation of Chinese characters) instead of writing characters. With regard to the finding in Chapter 5, it would be interesting to extend the electrophysiological measurement to test the lexico-syntactic encoding at the phrasal level and determine the temporal locus of lexico-syntactic encoding in speech production.

Moreover, Chapters 4 and 5 lend further evidence to the feasibility of investigating overt speech production with the electrophysiological measurement (Christoffels et al., 2007; Koester & Schiller, 2008; see Ganushchak et al., 2011 for a review). On the one hand, the electrophysiological measurement can provide fine-grained data to test the detailed time course of speech production. On the other hand, the correlation between he electrophysiological and behavioral data can provide a more solid reference for data interpretation.

To conclude, this dissertation provides empirical evidence for the understanding of the speech production processes and mechanisms in Mandarin Chinese, a language with a logographic script. It also contributes to the understanding of psycholinguistic models of speech production in general.

References

Aitchison, J. (2012). *Words in the mind: An introduction to the mental lexicon (4th ed.).* West Sussex, UK: John Wiley & Sons.

Alario, F.-X., & Caramazza, A. (2002). The production of determiners: Evidence from French. *Cognition, 82,* 179–223.

Alario, F. X., Perre, L., Castel, C., & Ziegler, J. C. (2007). The role of orthography in speech production revisited. *Cognition, 102,* 464-475.

Aristei, S., Melinger, A., & Rahman A. (2011). Electrophysiological chronometry of semantic context effects in language production. *Journal of Cognitive Neuroscience, 23,* 1567-1586.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390-412.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68,* 255-278.

Bates, E., Andonova, E., D'Amico, S., Jacobsen, T., Kohnert, K., Lu, C-C., Székely, A., … Pleh, C. (2000). Introducing the CRL International Picture-Naming Project (CRL-IPNP). *Center for Research in Language Newsletter, 12,* 1-14.

Bates, E., D'Amico, S., Jacobsen, T., Székely, A., Andonova, E., Devescovi, A., ... Tzeng, O. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review, 10,* 344-380.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R *package version 1.1-7.* http://CRAN.R-project.org/package= lme4.

Belke, E. (2008). Effects of working memory load on lexical-semantic encoding in language production. *Psychonomic Bulletin & Review, 15,* 357-363.

Belke, E. (2013). Long-lasting inhibitory semantic context effects on object naming are necessarily conceptually mediated: Implications for models of lexical-semantic encoding. *Journal of Memory and Language, 69,* 228-256.

Belke, E. (2017). The role of task-specific response strategies in blocked-cyclic naming. *Frontiers in Psychology, 7,* 1-6.

Belke, E., Meyer, A. S., & Damian, M. F. (2005). Refractory effects in picture naming as assessed in a semantic blocking paradigm. *The Quarterly Journal of Experimental Psychology, 58,* 667-692.

Belke, E., & Stielow, A. (2013). Cumulative and non-cumulative semantic interference in object naming: Evidence from blocked and continuous manipulations of semantic context. *The Quarterly Journal of Experimental Psychology, 66,* 2135-2160.

Bi, Y., Wei, T., Janssen, N., & Han, Z. (2009). The contribution of orthography to spoken word production: Evidence from Mandarin Chinese. *Psychonomic Bulletin & Review, 16,* 555-560.

Bi, Y., Xu, Y., & Caramazza, A. (2009). Orthographic and phonological effects in the picture–word interference paradigm: Evidence from a logographic language. *Applied Psycholinguistics, 30,* 637-658.

Bi, Y., Yu, X., Geng, J., & Alario, F.-X. (2010). The role of visual form in lexical access: Evidence from Chinese classifier production. *Cognition, 116,* 101–109.

Breining, B., Nozari, N., & Rapp, B. (2016). Does segmental overlap help or hurt? Evidence from blocked cyclic naming in spoken and written production. *Psychonomic Bulletin & Review, 23,* 500-506.

Caffarra, S., Janssen, N., & Barber, H. A. (2014). Two sides of gender: ERP evidence for the presence of two routes during gender agreement processing. *Neuropsychologia, 63,* 124–134.

Cai, Q. & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS ONE, 5,* e10729.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology, 14,* 177-208.

Caramazza, A., Miozzo, M., Costa, A., Schiller, N. O., & Alario, F.-X. (2001). A cross-linguistic investigation of determiner production. In E. Dupoux (ed.), *Language, Brain, and Cognitive Development: Essays in Honor of Jacques Mehler*, (pp. 209-226). Boston: MIT Press.

Chen, J. Y., Chen, T. M., & Dell, G. S. (2002). Word-form encoding in Mandarin Chinese as assessed by the implicit priming task. *Journal of Memory and Language, 46,* 751-781.

Cheng, L. L. S., & Sybesma, R. (1999). Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry, 30,* 509–542.

Cheng, L. L. S., & Sybesma, R. (2005). Classifiers in four varieties of Chinese. In G. Cinque & R. S. Kayne (eds.), *The Oxford Handbook of Comparative Syntax* (pp. 259–292). New York, NY: Oxford University Press.

Cheng, C. M. (1981). Perception of Chinese characters. *Acta Psychologica Taiwanica, 23,* 137–153.

Christoffels, I. K., Firk, C., & Schiller, N. O. (2007). Bilingual language control: An event-related brain potential study. *Brain Research, 1147,* 192-208.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82,* 407-428.

Costa, A., Strijkers, K., Martin, C., & Thierry, G. (2009). The time course of word retrieval revealed by event-related brain potentials during overt speech. *Proceedings of the National Academy of Sciences, 106,* 21442-21446.

Cubelli, R., Lotto, L., Paolieri, D., Girelli, M., & Job, R. (2005). Grammatical gender is selected in bare noun production: Evidence from the picture–word interference paradigm. *Journal of Memory and Language*, *53*, 42-59.

Damian, M. F. (2003). Articulatory duration in single-word speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 416-431.

Damian, M. F., & Als, L. C. (2005). Long-lasting semantic context effects in the spoken production of object names. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1372-1384.

Damian, M. F., & Bowers, J. S. (2003). Effects of orthography on speech production in a form-preparation paradigm. *Journal of Memory and Language, 49,* 119-132.

Damian, M. F., & Bowers, J. S. (2009). Assessing the role of orthography in speech perception and production: Evidence from picture–word interference tasks. *European Journal of Cognitive Psychology, 21,* 581-598.

Damian, M. F., Dorjee, D., & Stadthagen-Gonzalez, H. (2011). Long-term repetition priming in spoken and written word production: evidence for a contribution of phonology to handwriting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 813-826.

Damian, M. F., & Dumay, N. (2009). Exploring phonological encoding through repeated segments. *Language and Cognitive Processes, 24,* 685-712.

Damian, M. F., & Martin, R. C. (1999). Semantic and phonological codes interact in single word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 345-361.

Damian, M. F., & Stadthagen-Gonzalez, H. (2009). Advance planning of form properties in the written production of single and multiple words. *Language and Cognitive Processes, 24,* 555-579.

Damian, M. F., Vigliocco, G., & Levelt, W. J. (2001). Effects of semantic context in the naming of pictures and words. *Cognition, 81,* B77-B86.

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,* 124–142.

Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes, 5,* 313-349.

Dell, G. S., & O'Seaghdha, P. G. (1991). Mediated and convergent lexical priming in language production: A comment on Levelt et al (1991). *Psychological Review, 98,* 604-614.

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

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review, 104,* 801-838.

Dell'Acqua, R., Sessa, P., Peressotti, F., Mulatti, C., Navarrete, E., & Grainger, J. (2010). ERP evidence for ultra-fast semantic processing in the picture–word interference paradigm. *Frontiers in Psychology, 1,* 177.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134,* 9-21.

Doetjes, J. S. (1997). *Quantifiers and selection. On the distribution of quantifying expressions in French, Dutch and English*. Doctoral dissertation, Leiden University.

Donchin, E. (1981). Surprise!… surprise? *Psychophysiology, 18,* 493-513.

Duanmu, S. (2002). *The phonology of Standard Chinese.* Oxford: Oxford University Press.

Erbaugh, M. S. (1986). Taking stock: The development of Chinese noun classifiers historically and in young children. In C. Craig (ed.). *Noun Classes and Categorization* (pp. 399-436). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Fang, F. (1985). An experiment on the use of classifiers by 4-to 6-year-olds. *Acta Psychologica Sinica, 17,* 384-392.

Feldman, L. B., & Siok, W. W. T. (1999). Semantic radicals contribute to the visual identification of Chinese characters. *Journal of Memory and Language, 40,* 559-576.

Finkbeiner, M., & Caramazza, A. (2006). Now you see it, now you don't: On turning semantic interference into facilitation in a Stroop-like task. *Cortex, 42,* 790-796.

Finkbeiner, M., Gollan, T., & Caramazza, A. (2006). Bilingual lexical access: What's the (hard) problem? *Bilingualism: Language and Cognition, 9,* 153-166.

Fox, J., & Weisberg, S. (2011). *An R companion to applied regression* (2nd ed.). Thousand Oaks, CA: Sage.

Ganushchak, L., Christoffels, I., & Schiller, N. O. (2011). The use of electroencephalography in language production research: A review. *Frontiers in Psychology, 2,* 208.

Ganushchak, L., Verdonschot, R. G., & Schiller, N. O. (2011). When leaf becomes neuter: Event-related potential evidence for grammatical gender transfer in bilingualism. *Neuroreport, 22,* 106-110.

Glaser, W. R. (1992). Picture naming. *Cognition, 42,* 61-105.

Glaser, W. R., & Düngelhoff, F. J. (1984). The time course of picture-word interference. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 640-654.

Halekoh, U., & Højsgaard, S. (2014). A kenward-roger approximation and parametric bootstrap methods for tests in linear mixed models – the R package pbkrtest. *Journal of Statistical Software, 59,* 1-30.

Hothorn, T., Bretz, F. & Westfall, P. (2008). Simultaneous Inference in General Parametric Models. *Biometrical Journal, 50,* 346-363.

Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: Experimental and computational studies. *Cognition, 100,* 464-482.

Indefrey, P. (2011). The spatial and temporal signatures of word production components: A critical update. *Frontiers in Psychology, 2,* 255.

Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition, 92,* 101-144.

Janssen, N., Carreiras, M., & Barber, H. A. (2011). Electrophysiological effects of semantic context in picture and word naming. *NeuroImage, 57,* 1243–1250.

Janssen, N., Hernández-Cabrera, J. A., Van der Meij, M., & Barber, H. A. (2015). Tracking the time course of competition during word production: Evidence for a post-retrieval mechanism of conflict resolution. *Cerebral Cortex, 25,* 2960-2969.

Jescheniak, J. D., Hahne, A., & Schriefers, H. (2003). Information flow in the mental lexicon during speech planning: Evidence from event-related brain potentials. *Cognitive Brain Research, 15,* 261–276.

Jescheniak, J. D., Schriefers, H., Garrett, M. F., & Friederici, A. D. (2002). Exploring the activation of semantic and phonological codes during speech planning with event-related brain potentials. *Journal of Cognitive Neuroscience, 14,* 951–964.

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology and meaning* (pp. 67-84). Amsterdam, NL: Elsevier Science Publishers.

Koester, D., & Schiller, N. O. (2008). Morphological priming in overt language production: Electrophysiological evidence from Dutch. *NeuroImage, 42,* 1622-1630.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of event related potential (ERP). *Annual Review of Psychology, 62,* 621-647.

Kutas, M. & Van Petten, C. (1994). Psycholinguistics electrified: Event-related brain potential investigations. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 83-143). New York, NY: Academic Press.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). lmerTest: Tests in linear mixed effects models. *R package version 2.0-25.* http://CRAN.R-project.org/package=lmerTest.

La Heij, W. (1988). Components of Stroop-like interference in picture naming. *Memory & Cognition, 16,* 400–410.

La Heij, W., Mak, P., Sander, J., & Willeboordse, E. (1998). The gender-congruency effect in picture-word tasks. *Psychological Research*, *61*, 209-219.

Lau, E. F., Philips, C., & Poeppel, D. (2008). A cortical network for semantics: (de) constructing the N400. *Nature Reviews Neuroscience, 9,* 920-933.

Levelt, W. J. M (1992). Accessing words in speech production: Stages, processes and representations. *Cognition, 42,* 1-22.

Levelt, W. J. M. (1993). *Speaking: From intention to articulation.* Cambridge, MA: The MIT Press.

Levelt, W. J. M. (1999). Models of word production. *Trends in Cognitive Sciences, 3,* 223-232.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999a). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1-38.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999b). Multiple perspectives on word production. *Behavioral and Brain Sciences, 22,* 61-69.

Liu, Y., Hao, M., Li, P., & Shu, H. (2011). Timed picture naming norms for Mandarin Chinese. *PLoS ONE, 6,* e16505.

Luck, S. (2005). *An introduction to the event-related potential technique.* Cambridge, MA: The MIT Press.

Lupker, S. J. (1979). The semantic nature of response competition in the picture-word interference task. *Memory & Cognition, 7,* 485-495.

Lupker, S. J. (1982). The role of phonetic and orthographic similarity in picture–word interference. *Canadian Journal of Psychology/Revue Canadienne de Psychologie, 36,* 349-367.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163-203.

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007).

Lexical selection is not by competition: a reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 503-535.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG- data. *Journal of Neuroscience Methods, 164,* 177-190.

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

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

Miozzo, M., & Caramazza, A. (1999). The selection of determiners in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 907-922.

Miozzo, M., & Caramazza, A. (2003). When more is less: A counterintuitive effect of distractor frequency in picture-word interference paradigm. *Journal of Experimental Psychology: General, 132,* 228-252.

Miozzo, M., Costa, A., & Caramazza, A. (2002). The absence of a gender congruency effect in Romance languages: A matter of stimulus onset asynchrony? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 388-391.

Misra, M., & Holcomb, P. J. (2003). Event–related potential indices of masked repetition priming. *Psychophysiology, 40,* 115-130.

Mitchell, D. B., & Brown, A. S. (1988). Persistent repetition priming in picture naming and its dissociation from recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 14,* 213-222.

Navarrete, E., Del Prato, P., & Mahon, B. Z. (2012). Factors determining semantic facilitation and interference in the cyclic naming paradigm. *Frontiers in Psychology, 3,* 38.

Navarrete, E., Del Prato, P., Peressotti, F., & Mahon, B. Z. (2014). Lexical selection is not by competition: Evidence from the blocked naming paradigm. *Journal of Memory and Language, 76,* 253-272.

Navarrete, E., Mahon, B. Z., & Caramazza, A. (2010). The cumulative semantic cost does not reflect lexical selection by competition. *Acta Psychologica, 134,* 279-289.

Nickels, L., Biedermann, B., Fieder, N., & Schiller, N. O. (2015). The lexical-syntactic representation of number. *Language, Cognition and Neuroscience, 30,* 287–304.

O'Seaghdha, P. G., Chen, J. Y., & Chen, T. M. (2010). Proximate units in word production: Phonological encoding begins with syllables in Mandarin Chinese but with segments in English. *Cognition, 115,* 282-302.

Oostenveld, R., Fries, R., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience, 2011,* 156869.

Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2007). Cumulative semantic interference as learning. *Brain and Language, 103,* 175-176.

Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition, 114,* 227-252.

Peyraube, A. (1998). On the history of classifiers in Archaic and Medieval Chinese, in Tsou, B. K. (ed.), *Studia Linguistica Serica, Proceedings of the 3rd*

International Conference on Chinese Linguistics (pp. 39-68). Hong Kong: City University of Hong Kong Press.

Posnansky, C. J., & Rayner, K. (1978). Visual vs. phonemic contributions to the importance of the initial letter in word identification. *Bulletin of the Psychonomic Society, 11,* 188-190.

Praamstra, P., Meyer, A. S., & Levelt, W. J. (1994). Neurophysiological manifestations of phonological processing: Latency variation of a negative ERP component timelocked to phonological mismatch. *Journal of Cognitive Neuroscience, 6,* 204-219.

Protopapas, A. (2007). CheckVocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods 39,* 859–862.

Qu, Q., Damian, M. F., & Kazanina, N. (2012). Sound-sized segments are significant for Mandarin speakers. *Proceedings of the National Academy of Sciences, 109,* 14265-14270.

Qu, Q., Damian, M. F., & Li, X. (2016). Phonology contributes to writing: Evidence from a masked priming task. *Language, Cognition and Neuroscience, 31,* 251-264.

Rahman, A. R., & Melinger, A. (2009). Semantic context effects in language production: A swinging lexical network proposal and a review. *Language and Cognitive Processes, 24,* 713-734.

Rapp, B., & Caramazza, A. (2002). Selective difficulties with spoken nouns and written verbs: A single case study. *Journal of Neurolinguistics, 15,* 373-402.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition, 42,* 107-142.

Roelofs, A. (1993). Testing a non-decompositional theory of lemma retrieval in speaking: Retrieval of verbs. *Cognition, 47,* 59-87.

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, 64(3), 249-284.

Roelofs, A. (1999). Phonological segments and features as planning units in speech production. *Language and Cognitive Processes, 14,* 173-200.

Roelofs, A. (2000). WEAVER++ and other computational models of lemma retrieval and word-forming encoding. In L. Wheeldon (Ed.), *Aspects of Language Production* (pp. 71-114). East Sussex, UK: Psychology Press.

Roelofs, A. (2003). Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review, 110,* 88-125.

Roelofs, A. (2006). The influence of spelling on phonological encoding in word reading, object naming, and word generation. *Psychonomic Bulletin & Review, 13,* 33-37.

Roelofs, A. (2015). Modeling of phonological encoding in spoken word production: From Germanic languages to Mandarin Chinese and Janpanese. *Japanese Psychological Research, 57,* 22-37.

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

Rosinski, R. R., Golinkoff, R. M., & Kukish, K. S. (1975). Automatic semantic processing in a picture-word interference task. *Child Development, 46,* 247-253.

Rugg, M. D. (1985). The effects of semantic priming and word repetition on event-related potentials. *Psychophysiology, 22,* 642-647.

Rugg, M. D. (1990). Event-related brain potentials dissociate repetition effects of high-and low-frequency words. *Memory & Cognition, 18,* 367-379.

Schiller, N. O. (2007). Phonology and orthography in reading aloud. *Psychonomic Bulletin & Review, 14,* 460-465.

Schiller, N. O., & Caramazza, A. (2003). Grammatical feature selection in noun phrase production: Evidence from German and Dutch. *Journal of Memory and Language, 48,* 169–194.

Schiller, N. O., & Caramazza, A. (2006). Grammatical gender selection and the representation of morphemes: The production of Dutch diminutives. *Language and Cognitive Processes, 21,* 945–973.

Schmitt, B. M., Münte, T. F., & Kutas, M. (2000). Electrophysiological estimates of the time course of semantic and phonological encoding during implicit picture naming. *Psychophysiology, 37,* 473–484.

Schnur, T. T., Schwartz, M. F., Brecher, A., & Hodgson, C. (2006). Semantic interference during blocked-cyclic naming: Evidence from aphasia. *Journal of Memory and Language, 54,* 199-227.

Schnur, T. T., Schwartz, M. F., Kimberg, D. Y., Hirshorn, E., Coslett, H. B., & Thompson-Schill, S. L. (2009). Localizing interference during naming: convergent neuroimaging and neuropsychological evidence for the function of Broca's area. *Proceedings of the National Academy of Sciences, 106,* 322-327.

Schriefers, H. (1993). Syntactic processes in the production of noun phrases. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 841-850.

Schriefers, H., Meyer, A. S., & Levelt, W. J. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language, 29,* 86-102.

Schriefers, H., & Teruel, E. (2000). Grammatical gender in noun phrase production: The gender interference effect in German. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1368-1377.

Schuhmann, T., Schiller, N. O., Goebel, R., & Sack, A. T. (2009). The temporal characteristics of functional activation in Broca's area during overt picture naming. *Cortex, 45,* 1111-1116.

Schuhmann, T., Schiller, N. O., Goebel, R., & Sack, A. T. (2012). Speaking of which: Dissecting the neurocognitive network of language production in picture naming. *Cerebral Cortex, 22,* 701-709.

Severens, E., Van Lommel, S., Ratinckx, E., & Hartsuiker, R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. *Acta Psychologica, 119,* 159–187.

Shao, Z., Roelofs, A., Martin, R., & Meyer, A. S. (2015). Selective inhibition and naming performance in semantic blocking, picture-word interference, and color-word Stroop tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*, 1806-1820.

Siok, W. T., Perfetti, C. A., Jin, Z., & Tan, L. H. (2004). Biological abnormality of impaired reading is constrained by culture. *Nature, 43,* 71-76.

Snodgras, J. G. & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for naming agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 174-215.

Spalek, K., Damian, M. F., & Bölte, J. (2013). Is lexical selection in spoken word production competitive? Introduction to the special issue on lexical competition in language production. *Language and Cognitive Processes, 28,* 597-614.

Starreveld, P. A. (2000). On the interpretation of onsets of auditory context effects in word production. Journal of Memory and Language, 42, 497-525.

Starreveld, P. A., & La Heij, W. (1995). Semantic interference, orthographic facilitation, and their interaction in naming tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 686-698.

Starreveld, P. A., & La Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 896-918.

Starreveld, P., & La Heij, W. (2004). Phonological facilitation of grammatical gender retrieval. *Language and Cognitive Processes*, *19*, 677-711.

Strijkers, K., & Costa, A. (2011). Riding the lexical speedway: A critical review on the time course of lexical selection in speech production. *Frontiers in Psychology, 2,* 356.

Strijkers, K., Costa, A., & Thierry, G. (2010). Tracking lexical access in speech production: Electrophysiological correlates of word frequency and cognate effects. *Cerebral Cortex, 20,* 912-928.

Team, R. C. (2014). R: A language and environment for statistical computing. Vienna, Austria; 2014. URL http://www. R-project. org.

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comphrehension. *Proceedings of the National Academy of Science of the United States of America, 104,* 12530-12535.

Tsegaye, M. T, Mous, M., & Schiller, N. O. (2014). Plural as a value of Cushitic gender: Evidence from congruency effect experiments in Konso (Cushitic). In G. C. Corbett (Ed.), *The Expression of Gender* (pp. 191-214). Berlin/New York: Mouton de Gruyter.

Tzeng, O. J., Chen, S., & Hung, D. L. (1991). The classifier problem in Chinese aphasia. *Brain and Language*, *41*, 184-202.

Underwood, G, & P, Briggs (1984). The development of word recognition processes. *British Journal of Psychology, 75,* 243-255.

Van Berkum, J. J. (1997). Syntactic processes in speech production: The retrieval of grammatical gender. Cognition, 64, 115–152.

Van Casteren, M., & Davis, M.H. (2006). Mix, a program for pseudorandomization. *Behavior Research Methods, 38,* 584-589.

Van Turennout, M., Hagoort, P., & Brown, C. M. (1997). Electrophysiological evidence on the time course of semantic and phonological processes in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 787-806.

Van Turennout, M., Hagoort, P., & Brown, C. M. (1998). Brain activity during speaking: From syntax to phonology in 40 milliseconds. *Science, 280,* 572-574.

Verdonschot, R. (2011). *Word processing in languages using non-alphabetic scripts: The case of Japanese and Chinese.* Utrecht, NL: LOT.

Verdonschot, R., Lai, J., Chen, F., Tamaoka, K., & Schiller, N. O. (2015). Constructing initial phonology in Mandarin Chinese: Syllabic or subsyllabic? A masked priming investigation. *Japanese Psychologyical Research, 57,* 61-68.

Verdonschot, R. G., Nakayama, M., Zhang, Q., Tamaoka, K., & Schiller, N. O. (2013). The proximate phonological unit of Chinese-English bilinguals: Proficiency matters. *PLoS ONE, 8,* e61454.

Vigliocco, G., Vinson, D. P., Damian, M. F., & Levelt, W. (2002). Semantic distance effects on object and action naming. *Cognition, 85,* B61-B69.

Wang, L., Bastiaansen, M., & Yang, Y. (2015). ERP responses to person names as a measure of trait inference in person perception. *Social Neuroscience, 10,* 89–99.

Wang, L., Guo, J., Bi, Y., & Shu, H. (2006). Classifier congruency effect in the production of noun phrases. *Studies of Psychology and Behavior, 4,* 34-38.

Wheeldon, L. R., & Monsell, S. (1994). Inhibition of spoken word production by priming a semantic competitior. *Journal of Memory and Language, 33,* 332-356.

Wong, A. W. K., & Chen, H. C. (2008). Processing segmental and prosodic information in Cantonese word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 1172-1190.

Wong, A. W. K., & Chen, H. C. (2009). What are effective phonological units in Cantonese spoken word planning? *Psychonomic Bulletin & Review, 16,* 888-892.

Wu, Y., & Bodomo, A. (2009). Classifiers ≠ determiners. *Linguistic Inquiry, 40,* 487-503.

Yeh, S. L., & Li, J. L. (2004). Sublexical processing in visual recognition of Chinese characters: Evidence from repetition blindness for subcharacter components. *Brain and Language, 88,* 47-53.

Yoshihara, M., Nakayama, M., Verdonschot, R. G., & Hino, Y. (2017). The phonological unit of Japanese Kanji Compounds: A masked priming investigation. *Journal of Experimental Psychology: Human Perception and Performance,* http://dx.doi.org/10.1037/xhp0000374.

Yu, B., Feng, L., Cao, H., & Li, W. (1990). Visual perception of Chinese characters: Effect of perceptual task and Chinese character attributes. *Acta Psychologica Sinica, 22,* 141-148.

Zhang, J. J., & Liu, H. Y. (2009). The lexical access of individual classifiers in language production and comprehension. *Acta Psychologica Sinica, 7,* 580-593.

Zhang, Q., Chen, H. C., Weekes, B. S., & Yang, Y. (2009). Independent effects of orthographic and phonological facilitation on spoken word production in Mandarin. *Language and Speech, 52,* 113-126.

Zhang, Q., & Weekes, B. S. (2009). Orthographic facilitation effects on spoken word production: Evidence from Chinese. *Language and Cognitive Processes, 24,* 1082-1096.

Zhang, Q., & Yang, Y. (2003). The determiners of pictures - naming latency. *Acta Psychologica Sinica, 35,* 447–454.

Zhao, H., La Heij, W., & Schiller, N. O. (2012). Orthographic and phonological facilitation in speech production: new evidence from picture naming in Chinese. *Acta Psychologica, 139,* 272-280.

Zhou, L., Peng, G., Zheng, H. Y., Su, I. F., & Wang, W. S. (2013). Sub-lexical phonological and semantic processing of semantic radicals: a primed naming study. *Reading and Writing, 26,* 967-989.

Zhou, Y. G. (1978). Xiandai hanzihong shengpande biaoyin gongneng wenti [To what degree are the "phonetics" of present-day Chinese characters still phonetic?]. *Zhongguo Yuwen, 146,* 172–177.

Zhu, X., Damian, M. F., & Zhang, Q. (2015). Seriality of semantic and phonological processes during overt speech in Mandarin Chinese as revealed by event-related brain potentials. *Brain and Language*, *144*, 16-25.

Summary

This dissertation investigates the speech production of Mandarin Chinese from a psycholinguistic approach. Why is it interesting to investigate Mandarin Chinese speech production? From a theoretical point of view, current psycholinguistic models of speech production have been mainly based on evidence from West Germanic languages, where orthographic and phonological forms follow a certain mapping captured in grapheme-to-phoneme conversion (GPC) rules. By contrast, in languages with a logographic script such as Mandarin Chinese, GPC is more opaque, which may result in a (different) role for orthography in speech production. Previous research on the speech production of languages with a logographic script also provided empirical evidence suggesting possible modifications to the current speech production models.

Chapter 1 introduced the current psycholinguistic models of speech production. Most models agree that to overtly produce a word, speakers go through several stages: conceptual preparation, lemma retrieval, word-form encoding and articulation. At the word-form encoding stage, however, orthography and phonology are usually not dissociated. However, in languages with a logographic script like Mandarin Chinese, the orthographic representation of a lexical item - Chinese characters have a critical role in distinguishing homophones and may therefore be involved in speech production. In consequence, the speech production mechanisms of Mandarin Chinese may differ from what current models of speech production predict.

As the first experimental chapter, **Chapter 2** provided direct evidence for the involvement of orthography in speech production of Mandarin Chinese. Empirical evidence was reported to suggest the mandatory activation of orthography in speech production in English. More specifically, the spelling-sound inconsistency (e.g. joker, giant) will produce an inhibitory effect on

speech production, compared to consistency (e.g. joker, jewel). However, this finding was not replicated in Dutch, where presenting an orthographically-related but phonologically-unrelated word (e.g. cement) will not affect the naming latencies of the following target (e.g. congres).

There are at least two possible explanations for the discrepancy. One possibility is that the involvement of orthography may be task-dependent that only in tasks when the orthographic information is highly relevant, may there be the involvement of orthography in speech production. Another possibility is that the discrepancy may be attributed to the cross-linguistic differences. For instance, compared to that in English, the GPC is more transparent in Dutch, which may result in that orthography merely played any role in speech production in Dutch but some role in English.

In Chapter 2, we investigated the role of orthography in Mandarin Chinese using an adapted blocked cyclic naming paradigm. In this paradigm, participants were asked to overtly name pictures that were presented repeatedly in semantically homogeneous, phonologically homogeneous, or heterogeneous blocks. On each trial, a written Chinese character that was either orthographically related or unrelated to the target was briefly presented (for 75 ms) before the target picture. Consistent with previous research, an inhibitory semantic blocking effect and a facilitative phonological blocking effect in naming latencies were found. More importantly, we observed that the orthographically related characters facilitated picture naming in both the semantic and phonological blocks. In addition, the orthographic priming effect was independent of both the semantic and the phonological effects. These findings suggest orthography contributes to speaking in a picture naming task, lending further support to the presence of orthographic priming in spoken word production, at least in a language with a logographic script like Chinese.

The following chapter, i.e. **Chapter 3** investigated when and how orthography was involved during speech production. In previous research, the orthographic effect was observed at a similar stage to the semantic effect without the co-occurrence of any phonological effect. It was then suggested that orthography affected speech production via a lexico-semantic pathway. The critical evidence that supported this claim was that the orthographic effect was observed at negative SOAs (stimulus onset asynchrony) but this observation was not replicated in a later study. This chapter attempted to replicate it but did not observe any orthographic effect at negative SOAs in Experiments 1 or 2 and therefore suggested that it was unlikely that orthography affected speech production of Mandarin Chinese via a lexico-semantic pathway.

In Experiment 2 of Chapter 3, we used simplex characters to clearly dissociate orthography from the semantic representation and phonological representation. The orthographic effect was observed with the co-occurrence of the phonological effect, subsequent to the semantic effect. Since we made use of the simplex characters, i.e. characters without phonetic radical so that the GPC route was ruled out as a possible pathway. It is likely that orthographic relatedness affects speech production at another sub-lexical level, i.e. the character-to-syllable correspondence. More specifically, for a target (e.g. 兔, tu4, 'rabbit'), the orthographically related distractor (e.g. 免, mian3, 'exemption') activated its orthographic neighbors (e.g. 兔, tu4, 'rabbit'). Then, the activated character activated its syllable (tu4) and facilitated the speech production of the target.

Besides drawing evidence from behavioral data, in recent decades, researchers have increasingly used electrophysiological measurements to investigate the underlying mechanisms of speech production. With the high temporal resolution of electrophysiological measurements, Chapters 4 and 5 tapped into the neural correlates of speech production of Mandarin Chinese.

Chapter 4 investigated the neural correlates of semantic and phonological processing in speech production of Mandarin Chinese. Firstly, consistent with the findings in Chapter 2 and previous research, longer naming latencies were shown in semantically homogeneous blocks and shorter naming latencies in phonologically homogeneous blocks, relative to the heterogeneous blocks. Then, in the electrophysiological data, it was shown that the semantic factor significantly modulated electrophysiological waveforms from 200 ms and the phonological factor from 350 ms after picture presentation. The results are consistent with the estimation of meta-analyses on the neural correlates of speech production and studies using the go/no-go task and the picture-word interference task. In other words, the speech production of Mandarin Chinese also involves an earlier semantic processing and a later phonological processing and the temporal loci of these two stages are in line with those of the estimation of speech production in general.

Chapter 5 tapped into a more specific detail in the process of speech production; that is, whether a word's syntactic features (e.g. number, grammatical gender, etc.) are automatically activated and selected in bare noun production. Previous research has shown that the lexico-syntactic features are activated and selected in noun phrase production when these features are necessary for production. For instance, producing a noun phrase (e.g. de arm, 'the arm', *common gender*) will be facilitated by a distractor that is gender-congruent rather than incongruent (e.g. het been, 'the leg', *neuter gender*). However, it is debated if the lexico-syntactic features are activated and selected in bare noun production when these features are irrelevant for production. In Mandarin Chinese, although gender or case is not overtly marked, it is compulsory to use a classifier between a demonstrative and/or numeral and its associated noun.

Using the picture-word interference paradigm, we manipulated the congruency of Mandarin Chinese classifiers (i.e. a lexico-syntactic feature comparable to grammatical gender in psycholinguistic research) between the target picture (e.g. 'coat', *classifier-jian4*) and the superimposed distractor word (e.g. 'luggage', *classifier-jian4* or 'rabbit', *classifier-zhi1*). We measured the participants' naming latencies and their electroencephalogram (EEG). As a result, classifier incongruency elicited a stronger N400 effect in the ERP analyses, suggesting the automatic activation of lexico-syntactic features in bare noun production. However, classifier congruency did not affect naming latencies, suggesting that the lexico-syntactic feature is not selected in bare noun naming when it is irrelevant for production. It is possible that speech production in languages with relatively simple morphological structures, the selection at the lexico-syntactic layer is not necessary. The study of classifier effects provided insights to the comparison with regard to lexico-syntactic feature encoding between spoken word production in West-Germanic languages (where gender is a prominent feature) and that in East Asian languages (where classification is a prominent feature).

In summary, this dissertation investigated the speech production in Mandarin Chinese from a psycholinguistic perspective. The characteristic of opaque GPC in Chinese provides an interesting test case for the speech production, especially with regard of the separate roles of orthography and phonology. In this dissertation, it was shown that orthography contributed to speech production, probably by activating its orthographic neighbor and then the corresponding target syllable. In addition, pure orthographic relatedness could affect speech production independently without interacting with semantic or phonological representations. Moreover, this dissertation used electrophysiological measurement to investigate the fine-grained time course of speech production in Mandarin Chinese. It was shown that the semantic factor modulated the electrophysiological signals from 200 ms and the phonological factor from 350 ms after stimulus presentation. It was also shown that the lexico-syntactic feature (Chinese classifier) was automatically activated in speech production even when it was not necessary for production.

To conclude, this dissertation tapped into the details of speech production in Mandarin Chinese in the framework of current psycholinguistic models of speech production. The findings in this dissertation not only contribute to the understanding of the underlying neuropsychological mechanisms of speech production in Mandarin Chinese, but also provides insights into the understanding of the accountability of current models of speech production that are mostly based on evidence from West Germanic languages.

Samenvatting in het Nederlands

Deze dissertatie onderzoekt spraakproductie van het Mandarijn Chinees (MC) vanuit een psycholinguïstisch perspectief. Waarom het MC? Hedendaagse modellen van spraakproductie zijn voornamelijk gebaseerd op West-Germaanse talen zoals het Nederlands, het Engels en het Duits. Typisch voor deze talen is dat hun orthografische en fonologische vormen te beschrijven zijn in termen van grafeem-naar-foneem regels ("grapheme-to-phoneme conversion (GPC) rules" in het Engels). Zulke regels zijn in talen met zogenaamde logografische schriften zoals het MC nogal ondoorzichtig. Daarom zou orthografie in het MC een (andere) rol kunnen spelen in spraakproductie. Eerder onderzoek naar spraakproductie in logografische talen heeft inderdaad gewezen op mogelijke aanpassingen aan hedendaagse psycholiguïstische spraakproductiemodellen.

Hoofdstuk 1 introduceerde deze modellen, waarvan de meeste overeenkomen wat betreft de verschillende fases die doorlopen worden door een spreker tijdens het produceren van een woord, namelijk (1) conceptualisatie van de boodschap, (2) het vinden van het lemma, (3) het codificeren van de woordvorm en (4) de articulatie. Juist tijdens de derde fase verschillen de orthografie en fonologie weinig van elkaar. Echter, in logografische talen als het MC speelt de orthografische representatie van een lexicaal element een cruciale rol, zeker als het gaat om zogenaamde homofonen — woorden die hetzelfde klinken maar een verschillende betekenis hebben. Bijgevolg zouden de mechanismen die betrokken zijn bij spraakproductie in het MC kunnen verschillen van die in West-Germaanse talen.

Hoofdstuk 2 leverde experimenteel bewijs voor de rol van orthografie in spraakproductie in het MC. Eerst zijn bevindingen in studies in het Engels en Nederlands uiteengezet. In het Engels blijkt een verplichte activatie van orthografische representaties tijdens spraakproductie. Een tegenstrijdigheid in spelling en geluid (bijvoorbeeld "joker", "giant") leidt tot een belemmerend effect op spraakproductie, wanneer men dit vergeleek met consistente spelling-geluid paren (bijvoorbeeld "joker", "jewel"). Deze bevinding kon in het Nederlands niet gerepliceerd worden. Zo zal een orthografisch gerelateerd maar fonologisch ongerelateerd woord (bijvoorbeeld cement) geen effect hebben op het produceren van een volgend target woord (bijvoorbeeld congres).

Er zijn tenminste twee mogelijke verklaring voor de gevonden discrepantie. Het kan zijn dat de betrokkenheid van orthografie taak-afhankelijk is — alleen als orthografie relevant is speelt het een rol. Een andere mogelijkheid is dat de discrepantie voortvloeit uit cross-linguïstische verschillen. In het Nederlands zouden GPC regels bijvoorbeeld transparanter kunnen zijn vergeleken met het Engels. Die transparantie zou tot gevolg kunnen hebben dat orthografie nauwelijks een rol speelt in het Nederlands.

In het kader van deze bevindingen werd in hoofdstuk 2 de rol van orthografie in het MC onderzocht waarbij gebruik werd gemaakt van een "adapted blocked cyclic naming paradigm". Hierbij werden deelnemers gevraagd om plaatjes hardop te benoemen die herhaaldelijk in semantisch homogene, fonologisch homogene, of heterogene blokken gepresenteerd werden. Voorafgaand aan een target plaatje, verscheen er zeer kort (75 ms) een Chinees karakter dat orthografisch gerelateerd of ongerelateerd was (dit is ook wel bekend als "priming"). In lijn met eerder onderzoek kon een belemmerend semantisch effect ("blocking effect") en een faciliterend fonologisch blocking effect vastgesteld worden. Belangrijker nog, de orthografisch gerelateerde karakters hadden een faciliterend effect op de benoemingstaak in zowel de semantische als de fonologische blokken. Verder was het orthografische priming effect onafhankelijk van semantische en fonologische effecten. Deze bevindingen suggereren dat orthografie bijdraagt aan spraakproductie in een plaatjesbenoemingtaak. Het lijkt er op dat ook in een logografische taal als het MC orthografische priming vastgesteld kan worden.

In **Hoofdstuk 3** werd onderzocht wanneer en hoe orthografie betrokken is tijdens spraakproductie. Eerder onderzoek heeft aangetoond dat het orthografisch effect plaatsvindt vergelijkbaar met het moment dat een semantisch effect kan worden vastgesteld — zonder dat een fonologisch effect optreedt. Later heeft men voorgesteld dat orthografie via de lexicaal-semantische weg spraakproductie kan beïnvloeden. Doorslaggevend bewijs hiervoor was dat het effect van orthografie werd gemeten op negatieve "stimulus onset asynchronies" (SOAs). Dit kon evenwel niet worden gerepliceerd in een latere studie. Ook in dit hoofdstuk, waarin is geprobeerd orthografische effecten op negatieve SOAs in twee experimenten vast te stellen in het MC, zijn zulke effecten niet gevonden. Daarom lijkt het niet waarschijnlijk dat orthografie in de productie van het MC via de lexicaal-semantische route loopt.

In Experiment 2 in hoofdstuk 3 hebben we simplex MC karakters gebruikt om duidelijk de orthografie van de semantische en fonologische representaties te dissociëren. Na het semantisch effect kon zowel het orthografische effect als het fonologisch effect worden vastgesteld. Omdat we simplex karakters gebruikten, namelijk karakters zonder een fonetische radicaal zodat de GPC route kon worden uitgesloten, zou orthografische gerelateerdheid een ander sub-lexicaal niveau kunnen beïnvloeden — daar waar een karakter naar een lettergreep wordt omgezet. Meer specifiek, voor een target woord (e.g. 兔, tu4, 'rabbit'), activeerde de orthografisch gerelateerde afleider (e.g. 免, mian3, 'exemption') zijn orthografische buren (e.g. 兔, tu4, ‘rabbit’). Dus het geactiveerde karakter activeerde de corresponderende syllabe (tu4) and faciliteerde de spraakproductie van het target woord.

Naast de beschikbare (en almaar groeiende) gedragsdata hebben verschillende onderzoekers zich gericht op het gebruik van de elektrofysiologische techniek “event-related potentials” (ERPs) om de onderliggende mechanismen van spraakproductie te onderzoeken. Deze methode die bekend staat vanwege zijn hoge tijdsresolutie is gebruikt in hoofdstukken 4 en 5 om zo de neurale basis voor spraakproductie in het MC te achterhalen.

Hoofdstuk 4 onderzocht de neurale correlaten van semantische en fonologische verwerking van spraakproductie van MC. Tegelijkertijd werden gedragsdata verzameld die consistent bleken met de bevindingen als gerapporteerd in hoofdstuk 2 en eerder onderzoek. Vergeleken met heterogene blokken werd benoeming belemmerd in semantisch homogene blokken terwijl in fonologisch homogene blokken benoeming juist werd gefaciliteerd. In de elektrofysiologische data bleek de semantische factor een modulerend effect te hebben vanaf 200 ms, en de fonologische factor vanaf 350 ms — dit is gemeten na presentatie van een plaatje. Deze resultaten passen in de bevindingen van meta-analyses van neurale correlaten van spraakproductie en studies die gebruik maken van “go/no-go” en “picture-word interference” (PWI) taken. Met andere woorden, tijdens spraakproductie van het MC gaan semantische processen vooraf aan fonologische processen en het tijdsverloop is in lijn met spraakproductie in het algemeen.

Hoofdstuk 5 behandelde een meer specifiek onderdeel van het spraakproductieproces, namelijk, de syntactische kenmerken van een woord (bijvoorbeeld getal en grammatisch geslacht). In hoeverre worden zulke kenmerken automatisch geactiveerd en geselecteerd bij het benoemen van een zelfstandig naamwoord? Eerder onderzoek heeft aangetoond dat lexicaal-syntactische kenmerken van een woord worden geactiveerd zodra deze kenmerken nodig zijn voor de productie. Bijvoorbeeld, het uitspreken van de arm (gemeenschappelijk geslacht) zal worden gefaciliteerd door een afleider die congrueert qua geslacht maar niet door een woord met afwijkend geslacht zoals het been (neutraal geslacht). Het is de vraag of zulke kenmerken ook geactiveerd en geselecteerd worden als ze irrelevant zijn voor de productie. In het MC bestaat weliswaar geen markering voor geslacht, een zogenaamd klaswoord is wèl verplicht tussen een lidwoord, quantor of andersoortig modificerend element, en het corresponderende zelfstandige naamwoord. Klaswoorden worden daarom wel gezien als dragers van lexicaal-syntactische kenmerken.

Gebruikmakend van een PWI taak manipuleerden we in een ERP experiment de congruentie van klaswoorden in het MC van een target plaatje (bijvoorbeeld "coat", classifier-jian4) en het woord dat over het plaatje gepresenteerd werd (bijvoorbeeld "luggage", classifier-jian4 of "rabbit", classifier-zhi1). In de ERP analyses zagen we dat incongruentie van een klaswoord resulteerde in een sterker N400 effect. Het lijkt erop dat lexicaal-semantische kenmerken automatische geactiveerd worden bij het benoemen van een zelfstandig naamwoord. Echter, klaswoord congruentie had geen effect op de reactietijd van het benoemen. Blijkbaar worden lexicaal-semantische kenmerken niet geselecteerd tijdens het benoemen indien ze irrelevant zijn voor de productie. Mogelijk is de de selectie van lexicaal-semantische kenmerken in talen met relatief simpele morfologische structuren niet nodig. Het experiment heeft geholpen inzicht te brengen in hoeverre het codificeren van lexicaal-syntactische kenmerken in een Oost-Aziatische talen (waar klaswoorden een prominente rol vervullen) verschillen van West-Germaanse talen (waarin grammaticaal geslacht een prominente rol speelt).

Samenvattend heeft deze dissertatie spraakproductie van het MC vanuit een psycholinguïstisch perspectief onderzocht. Omdat GPC regels in een logografische taal als het MC niet voor de hand liggen is het een interessante casus voor spraakproductie, zeker in verband met de rollen die orthografie en fonologie spelen. Deze studie laat zien dat orthografie bijdraagt aan spraakproductie, waarschijnlijk door het activeren van orthografische buren en corresponderende lettergrepen. Louter orthografische gerelateerdheid lijkt spraakproductie te kunnen beïnvloeden zonder interactie van semantische en fonologische representaties. Verder is met behulp van ERPs het tijdsverloop van spraakproductie van het MC onderzocht. De semantische factor moduleerde de elektrofysiologische signalen vanaf 200 ms en de fonologische factor na 350 ms na presentatie van de stimulus. Bovendien is vastgesteld dat een lexicaal-syntactisch kenmerk in de vorm van een klaswoord automatisch geactiveerd wordt — zelfs als het niet nodig is voor productie.

Al met al dragen de resultaten van zowel de gedragsdata als de ERPs van spraakproductieprocessen in het MC bij aan hedendaagse psycholïnguistische modellen, die tot nog toe voornamelijk op basis van West-Germaanse talen ontwikkeld zijn.

Summary in Chinese 中文摘要

该博士论文从心理语言学的角度，研究汉语普通话的言语产出过程。为什么要研究汉语普通话的言语产出呢？从理论的角度看，现有的语言产出心理语言学模型主要建立于西方语言的基础上，尤其是英语、德语、荷兰语。这些语言有一个共同的显著特点，就是字音和字形的区分不够显著。也就是说，字形到字音的对应比较直接透明。相比较而言，在汉语等表意语言中，字形到字音的对应就没那么直接透明了。这种现象可能导致的后果是，字形在言语产出的过程中可能会扮演某种（与字母语言相比不同的）角色。已有研究发表了相关的实验结果，根据表意语言的言语产出实验数据，需要对现有的建立于字母语言基础上的语言产出模型进行修改。

本论文的**第一章**首先介绍了现有的语言产出的心理语言学模型。大多数的模型基本认可的是，说一个词，人们要经历概念上的加工，提取词条，字形加工，然后再把一个词说出来。然而，在字形加工（word-form encoding）这个阶段，模型里并没有仔细区分是字形 （orthography）还是字音（phonology）。在汉语中，一个汉字通常对应一个音节，然而汉语的音节数量是非常有限的，因此在区分同音字时，汉字扮演着至关重要的角色。那么，在汉语普通话的产出过程中，汉字到底有没有发挥作用呢？

作为第一个实验章节，**第二章**为回答这个问题提出了直接的证据。目前学术界还没有完全确定字形是否在言语产出中发挥作

用。英语的研究发现，相比于语音和拼写一致（joker, jewel；j 发音相同），如果语音相同拼写不一致的话（joker, giant），会延长言语产出的反应时间。但是在荷兰语和中文的相关实验里，并没有发现类似的结果。比如说，在荷兰语里，先呈现一个字形相关的词（cement），并不会加快说出接下来的目标词（congres）的反应速度（cement，congres 中 c 发音不同）。

造成这些分歧的可能原因至少有两个。一个是字形是否在言语产出中发挥作用，跟具体的实验任务有关。也就是说，当一个实验任务涉及字形的程度深的时候（如，阅读），字形才会对言语产出发生作用。另一个原因是跟具体的语言特征有关。比如说，荷兰语相比于英语，字形与字音的对应更直接透明，那么字形的作用可能就很小可以忽略不计；在英语中，字形则可能会发挥一定的作用。

在第二章中，我们采用了区组命名结合汉字启动的实验范式，让被试命名语义同质、语音同质以及异质区组里的图片，在每一张图片出现前，简短地呈现一个汉字（75 ms）。这个汉字与图片名称字形相关或无关。与文献里的发现一致，实验得到了区组命名里经典的语义干扰效应和语音协助效应。更重要的是，相比于字形无关的汉字，字形相关的汉字，显著加快了命名图片的反应时间。这一项结果表明，在汉语这样的表意语言里，字形在言语产出的过程中是发挥一定作用的。

接下来的问题就是，字形是在什么时刻，如何发挥作用的呢？在**第三章**中，我们开展了两个心理语言学实验，研究了这个问题。

在利用图字干扰范式的研究文献中，字形作用和语义作用出现在相同的 SOA（刺激呈现启动时差），而并没有伴随字音作用的出现。研究者认为，字形可能通过一个词－语义的通道协助词条的提取。但是这一发现并没有在后面的研究中得到认证。在第三章中，我们同样没有发现类似结果。相比较地，字形和字音的效应出现在相同的 SOA，与语义效应出现的 SOA 没有重合。因此，字形通过词－语义通道协助词条提取的假设，可能性比较小。

值得一提的是，在第三章的第二个实验里，我们只采用了独体字作为实验刺激。没有了声旁和义旁的干扰，字形和语义、语音可以更好地区分开来。实验二的结果基本复制了实验一的发现，也就是，字形和字音的效应出现在相同的 SOA。不含声旁的独体字让我们得以避免字形到字音的转换，那么字形发挥作用的路径就不是通过字形到字音对应的途径。讨论发现，字形（如，免）可能激活了字形上相似的汉字，其中包括目标词（如，兔），进而激活了目标音节（如，tu4），从而协助了语音加工，导致反应时加快。

除了收集行为数据以外，研究者越来越多地测量脑电活动，来研究言语产出的神经机制。本论文的第四章和第五章利用脑电数据提供的高时间分辨率，研究汉语普通话言语产出的神经机制。

第四章研究汉语普通话言语产出过程中的语义和语音加工的神经基础。我们采取了区组命名的范式，测量了反应时和行为数据。从脑电数据中可以发现，语义因子从图片呈现大约 200 ms 后开始显著影响脑电波形。相比较地，语音因子从图片呈现大约 350 ms 后开始显著影响脑电波形。这个结果和之前文献里用其他范式的西方语言以及汉语研究结果一致。这暗示着汉语普通话的语义和语音加工时刻与其他语言基本一致。

第五章研究了言语产出过程中更细微的一个环节，即词条的句法特征（如性、数、格）是否被激活和选择。文献里的研究发现，在命名名词词组的时候，词条的句法特征一致与否会对命名词组的反应时间长短产生影响。比如说，在目标图（如，胳膊，荷兰语 'de arm'，common gender）上呈现干扰词，相比于不同语法性别（如，腿，荷兰语 'het been'，neuter gender），相同语法性别会缩短命名目标图的反应时间。但是在仅名词的产出时，语法性别是否影响反应时长，在不同语言中的发现不一样。

在第五章中，我们采取图字干扰范式，设置了语义类别一致性和量词特征一致性两个因子，测量了被试的命名图片反应时长和脑电波活动。反应时长的数据呈现了经典的语义干扰效应，但是我们并没有发现量词效应，暗示着在仅名词本身的产出时，并不需要加工量词。有研究者提出语言产出有个必经的两层结构：词－语义层和词－句法层。在词形结构比较简单的语言中（比如汉语），句法特征并不需要被选择加工。但是量词特征是否会被

激活呢？我们发现量词一致性会影响脑电波，量词不一致会产生一个类似 N400 的脑电效应，暗示着量词（或其他句法特征）在仅名词的产出过程中，会被自动激活。量词的脑电效应的发现，为在现有言语产出的心理语言学模型框架下思考句法特征的加工提供了有价值的参考，也为比较西方语言，尤其西－日耳曼语言（以语法性别为显著特征）和东亚语言（以量词为显著特征）提供了更多的参考。

归纳起来，本博士论文从心理语言学的角度研究汉语普通话的产出。汉语字形和字音可以更好地分离，为研究语言产生中字形和字音各自的角色提供了很好的测试材料。我们通过一系列的行为和脑电实验发现，在汉语普通话的产出过程中，字形发挥一定的独立作用。字形可以通过激活相似汉字，从而激活目标音节，协助目标词的产出。本论文还发现在汉语普通话言语产出的过程中，在刺激呈现约 200 ms 后有显著的语义加工，约 350 ms 后有显著的语音加工。除此之外，汉语量词特征在仅名词的产出过程中，会被自动激活。

本博士论文在现有言语产生的心理语言学模型的框架下，研究汉语普通话的产出机制，不仅为理解汉语普通话的产出机制提供了有价值的实验证据，也为构建可以更广泛地解释语言现象的言语产出心理语言学模型，提供了有价值的参考。

Appendice

Appendix I Stimuli in Chapter 2: Semantic and phonological blocks.

Semantic blocks		
Semantica category	Target picture	Orthographically related primes
ANIMAL	骆驼 luo4tuo2 *camel*	驭 yu4 *drive*
	山羊 shan1yang2 *goat*	岁 sui4 *age*
	兔子 tu4zi0 *rabbit*	免 mian3 *exemption*
	天鹅 tian1e2 *swan*	夫 fu1 *husband*
CLOTHING	毛衣 mao2yi1 *sweater*	笔 bi3 *pen*
	衬衫 chen4shan1 *shirt*	村 cun1 *village*
	外套 wai4tao4 *coat*	夕 xi1 *sunset*
	裙子 qun2zi0 *skirt*	初 chu1 *beginning*

BODY PART	眼睛 yan3jing1 *eye*	退 tui4 *retreat*
	耳朵 er3duo0 *ear*	耷 da1 *drooping*
	胳膊 ge1bo0 *arm*	朕 zhen4 *I (used by the emporor)*
	肩膀 jian1bang3 *shoulder*	房 fang2 *house*
FURNITURE	衣柜 yi1gui4 *wardrobe*	袋 dai4 *bag*
	书桌 shu1zhuo1 *desk*	韦 wei2 *a family name*
	沙发 sha1fa1 *sofa*	吵 chao3 *noisy*
	凳子 deng4zi0 *stool*	凭 ping2 *on the basis of*

Phonological blocks		
Relatedness	Target picture	Orthographically related primes
bi	蝙蝠 bian3fu2 *bat*	虹 hong2 *rainbow*
	鼻子 bi1zi0 *nose*	自 zi4 *self*
	冰箱 bing1xiang1 *fridge*	泳 yong3 *swimming*
	别针 bie2zhen1 *safetypin*	另 ling4 *another*
ji	吉他 ji2ta1 *guitar*	壳 ke2 *shell*
	剪刀 jian3dao1 *scissors*	券 quan4 *coupon*
	镜子 jing4zi0 *mirror*	钩 gou1 *hook*
	金字塔 jin1zi4ta3 *pyramid*	全 quan2 *all*
qi	气球 qi4qiu2 *balloon*	氛 fen1 *atomosphere*
	跷跷板	挠

	qiao4qiao4ban3 *seesaw*	nao2 *scratching*
	青蛙 qing1wa1 *frog*	靛 dian4 *indigo*
	铅笔 qian1bi2 *pencil*	针 zhen1 *needle*
xi	香蕉 xiang1jiao1 *banana*	秃 tu1 *bald*
	仙人掌 xian1ren2zhang3 *cactus*	汕 shan4 *a place name*
	信封 xin4feng1 *envelop*	誉 yu4 *honor*
	犀牛 xi1niu2 *rhinoceros*	层 ceng2 *layer*

Appendix II Stimuli used in Chapter 3: Experiment 1 and 2.

Experiment 1				
	Distractor type			
	Semantically related		Semantically unrelated	
Target Picture	Orthograp-hically related	Orthographi-cally unrelated	Orthograp-hically related	Orthographi-cally unrelated
猩猩 xing1xing1 *gorilla*	狮 shi1 *lion*	鹅 e2 *goose*	独 du2 *alone*	柯 ke1 *a name*
吉他 ji2ta1 *guitar*	鼓 gu3 *drum*	琴 qin2 *piano*	喜 xi3 *favor*	知 zhi1 *knowledge*
桌 zhuo1zi0 *table*	床 chuang2 *bed*	窗 chuang1 *window*	杭 hang2 *a place name*	答 da2 *answer*
梨 li2 *pear*	杏 xing4 *apricot*	蕉 jiao1 *banana*	枪 qiang1 *gun*	缸 gang1 *jar*
椅子 yi3zi0 *chair*	柜 gui4 *closet*	凳 deng4 *stool*	构 gou4 *structure*	硫 liu2 *sulfur*
猫 mao1 *cat*	狗 gou3 *dog*	鹰 ying1 *owl*	犹 you2 *alike*	核 he2 *core*
碗 wan3 *bowl*	碟 die2 *plate*	盘 pan2 *plate*	矿 kuang4 *mine*	伯 bo2 *uncle*
胳膊 ge1bo0 *arm*	肚 du4 *belly*	头 tou2 *head*	服 fu2 *clothes*	权 quan2 *power*

腿 tui3 *leg*	脚 jiao3 *foot*	手 shou3 *hand*	朕 zhen4 *I (used by the emporor)*	钢 gang1 *steel*
花 hua1 *flower*	草 cao3 *grass*	叶 ye4 *leave*	艺 yi4 *art*	券 quan4 *coupon*
苹果 ping2guo3 *apple*	莓 mei2 *berry*	桔 ju2 *orange*	苍 cang1 *grey*	弧 hu2 *arc*
萝卜 luo2bo0 *radish*	葱 cong1 *onion*	姜 jiang1 *ginger*	节 jie2 *festival*	京 jing1 *a place name*
蘑菇 mo2gu1 *mushroom*	菜 cai4 *vegetable*	豆 dou4 *bean*	苏 su1 *a name*	库 ku4 *garage*
虾 xia1 *shrimp*	蜂 feng1 *bee*	鸡 ji1 *chicken*	虹 hong2 *rainbow*	福 fu2 *bless*
蜻蜓 qing1ting2 *dragonfly*	蛾 e2 *moth*	豹 bao4 *leopard*	蚀 shi2 *elipse*	模 mo2 *model*
蝎子 xie1zi0 *scorpion*	蟒 mang3 *python*	鸭 ya1 *duck*	褐 he4 *brown*	境 jing4 *place*
钉子 ding1zi0 *nail*	锤 chui2 *hammer*	斧 fu3 *axe*	钟 zhong1 *clock*	件 jian4 *piece*
锅 guo1 *pot*	铲 chan3 *spatula*	壶 hu2 *kettle*	铃 ling2 *bell*	地 di4 *ground*

饺子 jiao3zi0 *dumplings*	饼 bing3 *pastry*	面 mian4 *noodle*	馆 guan3 *place*	岛 dao3 *island*
骆驼 luo4tuo0 *camel*	驹 ju1 *horse*	鲑 gui1 *salmon*	骗 pian4 *lie*	坪 ping2 *grassland*

Experiment 2				
	Distractor type			
Target Picture	Semantically related	Phonologically related	Orthographically related	Unrelated
虫 chong2 *bug*	龟 gui1 *turtle*	充 **chong1** *charge*	**史** shi3 *history*	末 mo4 *end*
勺 shao2 *spoon*	叉 cha1 *fork*	少 **shao3** *few*	**句** ju4 *sentence*	川 chuan1 *river*
矛 mao2 *spear*	盾 dun4 *shield*	毛 **mao2** *fur*	**予** yu3 *I*	井 jing3 *well*
山 shan1 *mountain*	谷 gu3 *valley*	闪 **shan3** *blink*	**凶** xiong1 *bad luck*	瓦 wa3 *tile*
书 shu1 *book*	本 ben3 *(note)book*	术 **shu4** *skill*	**韦** wei2 *a family name*	月 yue4 *month*
牙 ya2 *tooth*	口 kou3 *mouth*	亚 **ya4** *Asia*	**才** cai2 *talent*	日 ri4 *sun*
鱼 yu2 *fish*	龙 long2 *dragon*	与 **yu3** *and*	**角** jiao3 *corner*	七 qi1 *seven*
尺 chi3 *ruler*	寸 cun4 *inch*	赤 **chi4** *red*	**户** hu4 *household*	辛 xin1 *a name*
虎 hu3 *tiger*	牛 niu2 *bull*	乎 **hu1** *a particle*	**虔** qian2 *sincere*	巾 jin1 *towel*

耳 er3 *ear*	头 tou2 *head*	儿 **er2** *son*	其 qi2 *its*	久 jiu3 *long*
石 shi2 *stone*	土 tu3 *sand*	式 **shi4** *pattern*	右 you4 *right*	六 liu4 *six*
目 mu4 *eye*	鼻 bi2 *nose*	母 **mu3** *mother*	且 qie3 *and*	文 wen2 *text*
刀 dao1 *knife*	匕 bi3 *dagger*	导 **dao3** *guide*	力 li4 *power*	卜 bu3 *a name*
风 feng1 *wind*	雨 yu3 *rain*	丰 **feng1** *a family name*	冈 gang1 *hill*	卢 lu2 *a family name*
人 ren2 *man*	工 gong1 *worker*	刃 **ren4** *knife edge*	八 ba1 *eight*	瓜 gua1 *melon*
手 shou3 *hand*	足 zu2 *foot*	兽 **shou4** *animal*	于 yu2 *at*	巴 ba1 *a name*
鼠 shu3 *mouse*	鸟 niao3 *bird*	束 **shu4** *bundle*	昆 kun1 *a name*	币 bi4 *money*
田 tian2 *farm*	农 nong2 *agriculture*	天 **tian1** *sky*	甲 jia3 *first*	气 qi4 *gas*
兔 tu4 *rabbit*	犬 quan3 *dog*	凸 **tu1** *convex*	免 mian3 *exemption*	厂 chang3 *factory*

羊	马	央	半	五
yang2	ma3	**yang1**	ban4	wu3
sheep	*horse*	*center*	*half*	*five*

Appendix III Stimuli used in Chapter 4: Semantic and phonological blocks

Semantic blocks			
Relatedness	Target picture	Relatedness	Target picture
ANIMAL	骆驼 luo4tuo0 *camel*	BODY PART	眼睛 yan3jing1 *eye*
	山羊 shan1yang2 *goat*		耳朵 er3duo0 *ear*
	兔子 tu4zi0 *rabbit*		胳膊 ge1bo0 *arm*
	天鹅 tian1e2 *swan*		肩膀 jian1bang3 *shoulder*
CLOTHING	毛衣 mao2yi1 *sweater*	FURNITURE	衣柜 yi1gui4 *wardrobe*
	衬衫 chen4shan1 *shirt*		书桌 shu1zhuo1 *desk*
	外套 wai4tao4 *coat*		沙发 sha1fa1 *sofa*
	裙子 qun2zi0 *skirt*		凳子 deng4zi0 *stool*

Phonological blocks			
bi	蝙蝠 bian3fu2 *bat*	*qi*	气球 qi4qiu2 *balloon*
	鼻子 bi1zi0 *nose*		跷跷板 qiao4qiao4ban3 *seesaw*
	冰箱 bing1xiang1 *fridge*		青蛙 qing1wa1 *frog*
	别针 bie2zhen1 *safetypin*		铅笔 qian1bi2 *pencil*
ji	吉他 ji2ta1 *guitar*	*xi*	香蕉 xiang1jiao1 *banana*
	剪刀 jian3dao1 *scissors*		仙人掌 xian1ren2zhang3 *cactus*
	镜子 jing4zi0 *mirror*		信封 xin4feng1 *envelop*
	金字塔 jin1zi4ta3 *pyramid*		犀牛 xi1niu2 *rhinoceros*

Appendix IV Stimuli used in Chapter 5.

		Distractor type			
		Semantically related		Semantically unrelated	
Target picture	Classifier	Classifier cong-ruent	Classifier incong-ruent	Classifier cong-ruent	Classifier incong-ruent
兔子 tu4zi0 *rabbit*	只 *zhi1*	企鹅 qi3e2 *penguin*	马 ma3 *horse*	袖子 xiu4zi0 *sleeve*	雨伞 yu3san3 *umbrella*
刀 dao1 *knife*	把 *ba3*	叉子 cha1zi0 *fork*	碗 wan3 *bowl*	扇子 shan4zi0 *hand fan*	雪茄 xue3jia2 *cigar*
裤子 ku4zi0 *pants*	条 *tiao2*	围巾 wei2jin1 *scarf*	雨衣 yu3yi1 *raincoat*	路 lu4 *road*	白菜 bai2cai4 *Chinese cabbage*
古琴 gu3qin2 *a musical instrument*	把 *ba3*	琵琶 pi2pa2 *a musical instrument*	大鼓 da4gu3 *a musical instrument*	火 huo3 *fire*	箭头 jian4tou2 *arrowhead*
叶子 ye4zi0 *leaf*	片 *pian4*	花瓣 hua1ban4 *petal*	树枝 shu4zhi1 *branch*	废墟 fei4xu1 *ruins*	夫妻 fu1qi1 *couple*
吉他 ji2ta1 *guitar*	把 *ba3*	二胡 er4hu2 *a musical instrument*	鼓 gu3 *drum*	斧子 fu3zi0 *axe*	毛笔 mao2bi3 *writing brush*

耳朵 er3duo0 *ear*	*只* *zhi1*	眼睛 yan3jing1 *eye*	头发 tou2fa4 *hair*	天鹅 tian1e2 *swan*	火车 huo3che1 *train*
小提琴 xiao3 ti2qin2 *violin*	*把* *ba3*	木琴 mu4qin2 *xylophone*	钢琴 gang1 qin2 *piano*	锁 suo3 *lock*	飞机 fei1ji1 *airplane*
手 shou3 *hand*	*只* *zhi1*	脚 jiao3 *foot*	头 tou2 *head*	鸭子 ya1zi0 *duck*	书桌 shu1 zhuo1 *desk*
手指 shou3zhi3 *finger*	*根* *gen1*	脚趾 jiao3zhi3 *toe*	指甲 zhi3jia3 *nail*	木头 mu4tou2 *wood*	奶酪 nai3lao4 *cheese*
支票 zhi1piao4 *check*	*张* *zhang1*	钞票 chao1 piao4 *bank note*	硬币 ying4bi4 *coin*	嘴 zui3 *mouth*	椅子 yi3zi0 *chair*
教堂 jiao4tang2 *church*	*座* *zuo4*	寺庙 si4miao4 *temple*	银行 yin2 hang2 *bank*	山 shan1 *mountain*	彩虹 cai3hong2 *rainbow*
松鼠 song1shu3 *squirrel*	*只* *zhi1*	猴子 hou2zi0 *monkey*	驴 lu:2 *donkey*	股票 gu3piao4 *stock*	眼镜 yan3jing4 *glasses*
壁虎 bi4hu3 *lizard*	*只* *zhi1*	章鱼 zhang1 yu2 *octopus*	公牛 gong1 niu2 *bull*	梨 li2 *pear*	词典 ci2dian3 *dictionary*

石头 shi2tou0 *stone*	块 *kuai4*	玉 yu4 *jade gemstone*	沙 sha1 *sand*	肉 rou4 *meat*	手套 shou3 tao4 *glove*
大衣 da4yi1 *coat*	件 *jian4*	毛衣 mao2yi1 *sweater*	帽子 mao4zi0 *hat*	行李 xing2li3 *luggage*	拖把 tuo1ba3 *mop*
蛋糕 dan4gao1 *cake*	块 *kuai4*	饼干 bing3gan1 *cookie*	冰淇淋 bing1qi1 lin2 *ice cream*	肌肉 ji1rou4 *muscle*	电脑 dian4 nao3 *computer*
卫生纸 wei4sheng1 zhi3 *toilet paper*	卷 *juan3*	画纸 hua4zhi3 *drawing paper*	餐巾纸 can1jin1 zhi3 *paper napkin*	胶卷 jiao1 juan3 *camera film*	萝卜 luo2bo0 *radish*
西红柿 xi1hong2shi4 *tomato*	个 *ge4*	柠檬 ning2 meng2 *lemon*	葱 cong1 *Welsh onion*	包 bao1 *bag*	墙 qiang2 *wall*
蛇 she2 *snake*	条 *tiao2*	龙 long2 *dragon*	猪 zhu1 *pig*	街 jie1 *street*	牙齿 ya2chi3 *teeth*
钢笔 gang1bi3 *fountain pen*	支 *zhi1*	铅笔 qian1bi3 *pencil*	尺子 chi3zi0 *ruler*	箭 jian4 *arrow*	钥匙 yao4shi0 *key*
照片 zhao4pian4 *photo*	张 *zhang 1*	相纸 xiang4 zhi3 *photographi c paper*	相机 xiang4ji1 *camera*	门票 men2 piao4 *entrance ticket*	足球 zu2qiu2 *football*

螃蟹 pang2xie4 *crab*	只 *zhi1*	虾 xia1 *shrimp*	鲤鱼 li3yu2 *common carp (type of fish)*	耳环 er3huan2 *earring*	镜子 jing4zi0 *mirror*
纸 zhi3 *paper*	张 *zhang1*	地图 di4tu2 *map*	笔 bi3 *pen*	床 chuang2 *bed*	汤 tang1 *soup*
老鼠 lao3shu3 *mouse*	只 *zhi1*	猫 mao1 *cat*	猛兽 meng3 shou4 *beast*	靴子 xue1zi0 *boot*	火柴 huo3chai2 *match*
河 he2 *river*	条 *tiao2*	小溪 xiao3xi1 *brook*	海 hai3 *sea*	毛巾 mao2jin1 *towel*	蚊子 wen2zi0 *mosquito*
牛 niu2 *cow*	头 *tou2*	狮子 shi1zi0 *lion*	鳗鱼 man3yu2 *eel*	大蒜 da4suan4 *garlic*	马车 ma3che1 *carriage*
香蕉 xiang1jiao1 *banana*	根 *gen1*	甘蔗 gan1zhe4 *sugar cane*	葡萄 pu2tao2 *grape*	汗毛 han4mao2 *fine hair*	灯塔 deng1ta3 *lighthouse*
衬衫 chen4shan1 *shirt*	件 *jian4*	衣服 yi1fu2 *clothes*	裙子 qun2zi0 *dress*	艺术品 yi4shu4pin3 *art piecework*	珠子 zhu1zi0 *bead*

Curriculum vitae

Man Wang was born in Shandong in the People's Republic of China in 1987. In 2009, she received her bachelor degree of Teaching Chinese as a Foreign Language in Shanghai International Studies University with a minor in International Finance in Fudan University (Shanghai). During 2007, she spent a semester in City University of Hong Kong as an exchange student to study Chinese linguistics. In 2009, she received a full scholarship and joined the Computational Linguistics Lab (now the Lab of Computational Linguistics & Brain and Cognitive Neuroscience) in Tsinghua University (Beijing) and received her master degree in 2012. She then started working at Leiden University Center for Linguistics as a PhD researcher. The dissertation is the result of her research.