The lexico-semantic representation of words in the mental lexicon = De lexico-semantische representatie van woorden in het mentale lexicon

Wang, Y.

CHAPTER 1

General Introduction

Speech production is a fundamental aspect of human communication, enabling us to convey meaning and interact with others effortlessly. Although it appears effortless, the cognitive architecture underlying speech production is complex, involving the orchestration of various systems, including multiple cognitive and physiological components. These include retrieving information from memory, arranging words in a proper grammatical structure and order, and articulating words into syllables.

In this thesis, we will focus on the cognitive mechanisms of speech production. We will explore how the brain processes linguistic information. By examining these aspects, we aim to uncover the underlying architecture that makes speech production possible and understand the complexities involved in this remarkable human ability.

1.1 Levelt's model of word production

Speech production refers to the procedure of transforming thoughts into spoken utterances. This procedure involves retrieving words from memory that convey the intended meaning, formatting them appropriately (e.g., adjusting for plural, case, etc.), arranging them in the correct order before verbalizing the utterance. To explain word production, multiple models of word production have been developed. For instance, interactive models (e.g., Dell, 1986, 2013), cascaded models (e.g., Caramazza, 1997), and discrete or serial models of language production (Levelt, 1999; Levelt et al., 1999).

Levelt's model assumes that language production consists of three sequential strata: the conceptual/semantic stratum, the lemma stratum, and the phonological word-form stratum(Levelt, 1999; Levelt et al., 1999). At the conceptual stratum, Levelt's model assumes that concepts can be represented by holistic nodes connected to associated hypernyms/semantic categories in a semantic network. When retrieving words for word production, a preverbal concept first needs to be activated sufficiently. Importantly, at this early stage of word production, multiple concepts that are sufficiently close in terms of meaning to the intended concept are also activated to a sufficiently large degree. The preverbally intended and other unintended, but sufficiently activated concepts then activate their corresponding lemmas at the second level, i.e., the lemma stratum. By activating their corresponding lemmas, those concepts become lexicalized, i.e., the given concepts will be transformed into modality-free, lexical representations, so-called lemmas.

The lemma of a lexical concept has a unidirectional flow of information to its lexico-syntactic features (Levelt, 1999; Levelt et al., 1999; Roelofs, 1992). Through this unidirectional information flow, the corresponding, grammatically correct options of lexico-syntactic features (e.g., gender, number, case, etc.) are automatically activated. Multiple activated lemmas will compete for selection (lemma competition), with only one of the lexicalized concepts eventually being selected. This single selected lemma will then activate its corresponding phonological word form at the phonological stratum, allowing for it to be articulated (Levelt, 1999; Levelt et al., 1999; Roelofs, 1992, 1996). Note that other models (e.g., interactive and cascaded models of language production) assume that phonological word forms of activated but unselected lemmas can also become activated (Dell, 2013; Navarrete & Costa, 2005; Peterson & Savoy, 1998; see overview in Schiller & Alario, 2023).

1.2 Limitations of Levelt's model

Although Levelt's model is generally considered to be one of the most comprehensive theoretical models of language production, it has its limitations. First, at the conceptual stratum, Levelt's model theorizes that it is the hypernyms/semantic categories (e.g., animals, furniture, etc.) that determine the construction of a semantic network. However, hypernyms/semantic categories are typically correlated with some semantic features as they are defined as collections of concepts that share multiple more elementary semantic features such as shape and animacy (Hutchison, 2003; McRae et al., 1997). These semantic features are defined by Reilly et al. as "... *component[s] or element[s] that relate to a concept or express[es] [their] relation[s] with other concepts. Semantic features capture a wide range of information characteristics of a concept covering taxonomic relations, perceptual properties, function, behavior, thematic roles, and introspective features*" (Reilly et al., 2024). The inherent correlation of semantic categories and semantic features makes it uncertain if the hypernyms/semantic categories themselves determine the construction of a semantic network or it is instead driven by the (associated) semantic features. Hence, the question arises which role semantic features play in word production.

Second, although Levelt's model theorizes that semantic categories (hypernyms) play a pivotal role in the organization of the semantic network, the role of semantic features is less clear in their model. To illustrate, both *mammal* and *animal* are semantic categories (hypernyms) for *elephant* and *cow*. Furthermore, in the case of two concepts sharing multiple semantic categories, these semantic categories are hierarchically ordered. E.g., in the previous example, *mammal* is a more specific descriptor than *animal*. This hierarchical ordering of semantic categories (hypernyms) for a given word-pair might correlate with the difference in the amount of shared semantic features between the higher-ranking semantic category vs. the lower-ranking semantic category. For example, compared to *ele-*

phant and *bird*, *elephant* and *cow* have more overlapping features (e.g., having a tail, having four legs, etc.). However, this case was not considered in Levelt's model. Hence, it led to the question if semantic features play a role in word production, and what role the number of overlapping features play in word production.

Last, Levelt's model has been developed based on experimental findings of studies that assumed that a word's lexico-syntactic features (can) only have a single grammatically correct option (Huang & Schiller, 2021; Levelt, 1999; Levelt et al., 1999; Schiller & Caramazza, 2003; Schriefers, 1993; Wang et al., 2019). Given the languages and types of lexico-syntactic features used in those studies, this assumption was appropriate. For example, most words in Dutch and German have only one grammatical gender (for dialectal variations in German, see Schiller, 2025). However, this assumption is overly simplistic for studies investigating other languages, such as e.g. Mandarin Chinese, as they contain lexico-syntactic features which allow for multiple grammatically correct options for these features. To illustrate, in Mandarin Chinese, nouns typically have multiple grammatically correct options regarding classifiers (i.e., lexico-syntactic features in Mandarin Chinese) (Liu et al., 2019; Wu & Bodomo, 2009). However, to date, researchers that studied Mandarin Chinese classifiers have simplified that for a given noun, only the most compatible classifiers (i.e., dominant classifier) was correct (Huang & Schiller, 2021; Wang et al., 2019). Therefore, we are interested in understanding the effect of multiple grammatically correct options for a lexico-syntactic feature in language production in the context of Levelt's model. In the case of Mandarin Chinese specifically, the question becomes how classifiers are activated in word production.

To address these limitations of Levelt's model, we explore the following central themes in this thesis. First, we investigate if single and multiple semantic features play a role in speech production. Second, we study if the lemma stratum as conceptualized in Levelt's model can account for the processing of lexico-syntactic features of nouns when multiple grammatically correct options are possible.

1.3 The experimental framework in this thesis

In this thesis, we employed the picture–word interference (PWI) paradigm (Glaser, 1992; Lupker, 1979; Rosinski et al., 1975) together with electrophysiological measurements (i.e., electroencephalography; EEG) to explore the described limitations of Levelt's model through investigating the themes discussed in the previous paragraph. PWI is a classic paradigm to investigate the underlying cognitive mechanisms regarding linguistic features (e.g., semantic, syntactic, phonological, etc.) in word production.

In the PWI paradigm, participants are presented with a target picture superimposed with a written distractor word and asked to overtly name the target while ignoring the superimposed written distractor word. The research question is answered by carefully manipulating the linguistic relationship between the target and the distractor word. In **Chapters 2** and **3**, regarding the role of single and multiple semantic features in word production, we manipulated the congruency of single and/or multiple semantic features. In **Chapter 4**, where we investigate the processing of lexico-syntactic features of nouns when multiple grammatically correct options are possible, we first represented options of a lexico-syntactic feature (i.e., Mandarin Chinese classifier) of a noun as a probability distribution. Then we manipulated Mandarin Chinese classifier probability similarity between target and distractor words in PWI, akin to the congruency of grammatical gender in Indo-European languages. Through such manipulation, we can investigate its effect on naming accuracies and naming latencies.

In addition to the standard behavioural measures obtained through the PWI paradigm, we employed EEG to obtain time-course information during language production. EEG is a non-invasive method to measure event-related brain potentials (ERPs), which are large-scale systematic oscillatory changes of brain potentials over time upon encountering a given event, such as, e.g., the onset of a stimulus during a task (Woodman, 2010). In the studies of language production, the N400 and P600 components are often investigated. The N400 component is a negative deflection primarily centered over the centro-parietal regions and observed between 250 and 500 ms after stimulus onset. It exhibits a maximum at approximately 400 ms post-stimulus onset (Kutas & Hillyard, 1980a, 1980b, 1984). In language production, the N400 component has been identified as a marker of the process where speakers select words that fit the discourse context, indicating semantic integration. In language production, more negative N400 effects have been observed in scenarios such as naming objects in categorically unrelated versus related contexts in object naming tasks (e.g., PWI) (Blackford et al., 2012; Greenham et al., 2000; Wang et al., 2019). This N400 effect is attributed to lemma-level competition (Costa et al., 2009; Wang et al., 2019).

In contrast, the P600 component is a positive-going deflection centered around centro-parietal regions, having an onset around 500 ms post-stimulus onset and peaking around 600 ms post-stimulus onset (Osterhout & Holcomb, 1992). It has been proclaimed as an indicator of syntactic processing (Hagoort & Brown, 2000; Popov et al., 2020). In summary, the N400 effect is driven by semantic factors, such as semantic categories, and/or semantic features, whilst the P600 effect is driven by syntactic factors, such as the syntactic elements of classifiers (Blackford et al., 2012; Costa et al., 2009; Greenham et al., 2000; Wang et al., 2019; for an overview, see Wang & Schiller, in press). In **Chapters 2** and **3** of this thesis, semantic aspects of words are manipulated and hence an N400 is expected. In **Chapter 4**, lexico-syntactic aspects of words are manipulated and hence a P600 is expected.

1.4 The current thesis

In **Chapter 2** of this thesis, we examine the role of a particular semantic feature (i.e., animacy) in language production in a PWI study conducted in Mandarin Chinese whilst controlling for classifier congruency and recording behavioural and electrophysiological responses. We observed an animacy interference effect together with a larger N400 component for animacy-incongruent versus congruent picture-word pairs, suggesting that the semantic feature animacy can pre-activate concepts sharing this feature.

In **Chapter 3**, we investigate the number of overlapping semantic features in language production in a PWI study similarly conducted in Mandarin, varying semantic category and shape congruency whilst controlling for classifier congruency, recording behavioural and electrophysiological responses. We observed a main effect of semantic category that was stronger than the main effect of shape (i.e., a semantic feature). That is, the reduction in naming accuracies, the increase in naming latencies, and the increase in ERP amplitudes between 275 – 575 ms post-stimulus onset (N400 effect) for congruent vs. incongruent conditions were larger for semantic category than shape. In addition, we found an interaction effect between semantic category and the semantic feature 'shape' regarding naming accuracies and voltage amplitudes. From our results we conclude that, with an increase in the number of overlapping features between word pairs, there may be more spreading of activation between them during word production.

In **Chapter 4**, we asked native Mandarin Chinese speakers to name target pictures using the PWI paradigm while recording EEG to track the time course more precisely. Distractor words with varying degrees of classifier distribution similarity were superimposed while controlling for dominant classifier congruency. Distractors with dissimilar classifier distributions resulted in a more positive P600-like effect, but no behavioural effect was observed compared to distractor nouns having similar classifier distributions. Based on this result, we propose that, when producing a bare noun, multiple compatible classifiers are activated at the lexical level.

Finally, **Chapter 5** serves the purpose of integrating the findings from each of the studies reported in this thesis. Therefore, in this chapter, we synthesize our research findings and their theoretical implications, while also discussing future directions.